THE BRITISH JOURNAL OF STATISTICAL PSYCHOLOGY

EDITED BY
CYRIL BURT AND JOHN WHITFIELD

WITH THE ASSISTANCE OF
CHARLOTTE BANKS AND ALAN STUART
AND THE FOLLOWING EDITORIAL BOARD

A. C. AITKEN
M. S. BARTLETT
W. G. EMMETT
M. G. KENDALL
D. N. LAWLEY
E. S. PEARSON
E. A. PEEL
L. S. PENROSE
J. FRASER ROBERTS
A. RODGER
W. STEPHENSON
P. E. VERNON

VOLUME XI

Printed and Published by
TAYLOR & FRANCIS LTD.
RED LION COURT, FLEET STREET, LONDON, E.C.4

Digitized by the Internet Archive
in 2026

https://archive.org/details/british-journal-of-mathematical-statistical-psychology_1958-05_11_part-1

Vol. XI Part I The British Journal of Statistical Psychology May 1958

ESTIMATION IN FACTOR ANALYSIS UNDER VARIOUS INITIAL ASSUMPTIONS

By D. N. Lawley
University of Edinburgh

The object of the following paper is to investigate how far it is practicable to adopt the principle of maximum likelihood for the estimation of factor loadings, when some of the loadings are assumed *a priori* to be zero. Both the orthogonal and the oblique case are considered; and the solutions obtained are illustrated by applying them to a simple example.

Owing to the need for an iterative procedure the labour involved increases enormously as the number of factors increases. It is concluded that one possibility would be to use methods which, though not fully efficient, would entail only fairly simple calculations. To devise such methods is comparatively easy, but it is far more difficult to assess their efficiency.

I. Introduction

It is well known in factor analysis that an assumption regarding the number of factors present (supposing this to be greater than one) is not in itself sufficient to enable these factors to be uniquely determined from the covariance or correlation matrix of the test variates. In practice rotation is usually carried out in order to reach a solution which is satisfactory from a psychological point of view. In some cases a certain number of zero loadings are postulated in order to obtain a unique solution. A well known example of this is the idea of 'simple structure' put forward by Thurstone (see, for example, Thurstone [6]). One great difficulty of this method from a statistical standpoint is that the exact number and location of the zeros are not, as a rule, specified beforehand, but are decided only after an examination of the data. In certain cases, however, it seems likely that the experimenter could postulate in advance that certain specified loadings are zero. If these were sufficient in number and suitably placed, that would completely determine the factors and the set of loadings. The main problem then would be that of providing a satisfactory set of estimates of these loadings. Another question which arises is whether the factors are assumed to be orthogonal or oblique. In the latter case it will clearly be necessary to estimate the correlations between the factors.

These and other questions have been discussed in some detail in two recent studies, by Anderson and Rubin [1] and by Howe [4]. The latter in particular has devoted much valuable research to the practical solution of the various maximum likelihood estimation equations which arise; and his contribution is strongly recommended to anyone interested in this aspect.

Unaware of these researches the writer independently began an investigation of the possibilities of applying the maximum likelihood method to the estimation

of factor loadings when certain of them are assumed *a priori* to be zero, both for the orthogonal and for the oblique case. Certain data, previously analysed by Emmett [3] by means of the ordinary maximum likelihood method, were used to illustrate the practical solution of the equations of estimation. In view of the work of the authors already mentioned the work has lost some of its originality. It is presented here, however, in the hope that further discussion may still be useful, especially since the previous researches may not be familiar or readily available.

II. Estimation with Orthogonal Factors

(*a*) *Equations of Estimation.* Suppose that p normally distributed test variates x_i $(i = 1, 2, ..., p)$ are assumed to depend upon k common factors f_r $(r = 1, 2, ..., k)$ and that l_{ir} is the true value of the loading of x_i in f_r. We shall suppose first that the f_r are orthogonal, i.e. uncorrelated, with unit variances. Then if C is the true covariance matrix of the x_i we may write

$$C = LL' + V,$$

where $L = [l_{ir}]$ is the matrix of loadings and where V is the diagonal matrix whose elements consist of the residual variances σ_i^2 of the x_i. These residual variances represent the combined effects of specific factors and error.

In accordance with our previous remarks certain of the l_{ir} will be assumed *a priori* to be zero. These will in future be referred to simply as 'zero loadings'.

Let A be a sample covariance matrix whose elements a_{ij} are the usual estimates of the variances and covariances of the x_i with n degrees of freedom (corresponding as a rule to a sample of size $n+1$). Then, since the elements of A follow a Wishart distribution, the log-likelihood function may be written (omitting a function of the observations) as

$$-\tfrac{1}{2}n \log_e |C| - \tfrac{1}{2}n \operatorname{tr}(AC^{-1}), \tag{1}$$

where tr denotes the trace or sum of diagonal elements.

Maximum likelihood estimates are obtained by maximizing this expression with respect to the non-zero l_{ir} and with respect to the σ_i^2, assuming these to be unknown. If estimates are denoted by circumflex accents and if $\hat{C} = \hat{L}\hat{L}' + \hat{V}$, we obtain, after a certain amount of algebra, two sets of equations. The first may be expressed by the statement that those elements of the matrix

$$Q' = \hat{L}'\hat{C}^{-1} - \hat{L}'\hat{C}^{-1}A\hat{C}^{-1} \tag{2}$$

which correspond in position to the non-zero elements of L' (and of $\hat{L}'$) are to be equated to zero. These equations were in fact given earlier by the writer (Lawley [5]). The other set of equations may be written as

$$\operatorname{diag}(\hat{C}^{-1} - \hat{C}^{-1}A\hat{C}^{-1}) = 0, \tag{3}$$

where $\operatorname{diag}(X)$ means the matrix consisting only of the diagonal part of X.

These equations are not in forms which are suitable for practical solution, but they may be transformed to make them more so. It will be convenient

at this point to discard all circumflex accents. By L, V and C we shall in future mean $\hat{L}$, $\hat{V}$ and $\hat{C}$. If we post-multiply by V in (2) and remember that $V = C - LL'$, we obtain

$$Q'V = L' - L'C^{-1}A - KL', \tag{4}$$

where K is a symmetrical $k \times k$ matrix given by

$$K = L'C^{-1}(C - A)C^{-1}L = Q'L.$$

Since V is diagonal, $Q'V$ has zeros in the same positions as Q'. It may be observed that, in view of the properties of Q', the diagonal elements of K must be zero. The non-diagonal elements of K are not, however, necessarily zero; on this more will be said later.

To simplify (3) we first post-multiply the matrix on the left-hand side by V, in the form $C - LL'$. The diagonal elements are still zero; and, in view of the fact that $\text{diag}(QL') = 0$, we find that

$$\text{diag}(I - C^{-1}A) = 0,$$

where I denotes the unit matrix, in this case of order p. Pre-multiplication by V then yields the equation

$$V = \text{diag}(A - LL'C^{-1}A). \tag{5}$$

Except in particular cases this does not lead to the usual estimates for the residual variances; and the diagonal elements of C are not, in general, equal to the corresponding elements of A. Equations (4) and (5), though expressed in different notation, are the same as those reached by Howe; and the following discussion of their practical solution has much in common with his, though some points are here given greater emphasis.

As an example let us consider seven variates depending on three common factors. Using x's to denote the non-zero loadings, let us suppose that the pattern of these is assumed to be as follows:

x	x	x	x	x	x	x
x	x	x	x	x	0	0
0	0	0	x	x	x	x

It will be noticed that the tests having non-zero loadings in the second factor form a sub-set of those having non-zero loadings in the first factor. A brief examination of the positions of the zeros in L and Q is sufficient to show that, as a consequence, $k_{12} = k_{21} = 0$, where k_{rs} is the typical element of K. Similar considerations apply to the first and third factors, and so also $k_{13} = k_{31} = 0$. When, however, the second and third factors are paired, the set of tests having non-zero loadings in either one is not a sub-set of those having non-zero loadings in the other. Hence $k_{23} = k_{32}$ is not necessarily zero. This is the only non-zero element of K.

Now denote the three rows of L', in order, by l'_1, l'_2 and l'_3. Then, in this example, (4) implies that the non-zero elements l'_1, l'_2 and l'_3 are equal to the

corresponding elements of the row vectors u_1', u_2' and u_3', respectively, where

$$u_1' = l_1' C^{-1} A,$$
$$u_2' = l_2' C^{-1} A + k_{23} l_3',$$
$$u_3' = l_3' C^{-1} A + k_{23} l_2'.$$

This, as will be seen, leads to an iterative method for solving the equations of estimation. For purposes of computation we note that the inverse of C is given by

$$C^{-1} = V^{-1} - V^{-1} L (I+J)^{-1} L' V^{-1},$$

and that

$$L' C^{-1} = (I+J)^{-1} L' V^{-1},$$

where $J = L' V^{-1} L$. The calculations required are more tedious than those that occur in the ordinary maximum likelihood method since in each iteration the inversion of the $k \times k$ matrix $I+J$ is necessary; and this is not in general diagonal.

Suppose instead that the pattern of loadings were as shown below:

x	x	x	x	x	x	x
x	x	x	x	0	0	0
0	0	0	0	x	x	x

Then, as before, k_{23} is the only non-zero element of K; but now contributions arising from k_{23} may be ignored. The reason is that there is no 'overlap' between the loadings in the second and third factors, i.e. there are no tests having non-zero loadings in both factors. In this case the non-zero elements of L' are equal to the corresponding elements of $L' C^{-1} A$.

Thus for certain patterns of loadings the equations arising from (4) can be simplified by omitting the term in K, in which case (5) can be simplified to $V = \mathrm{diag}(A - LL')$, giving the usual estimates of the residual variances. This simplification is possible, for example, with the so-called 'bifactor' type of analysis. For such cases Howe recommends a Gauss–Seidel type of iterative technique[1] which will not, however, be described here. The method is very laborious; but he claims that it gives more rapid convergence than other methods. It may be noted that for both the patterns given above the loadings are uniquely determined, since in either case any orthogonal rotation of the factors would destroy the arrangement of zeros.

(*b*) *Iterative Method of Solution.* In the numerical example to be used for illustrating the method of solution there were eight tests, depending on three factors; and the pattern of loadings was assumed to be:

x	x	x	x	x	x	x	x
x	x	x	0	0	0	0	0
0	x	x	x	x	x	x	0

[1] Howe also recommends this technique for the ordinary maximum likelihood method (where there are no *a priori* zero loadings) in cases where an exact solution of the equations of estimation is desired.

It must be made clear that this is not really an *a priori* hypothesis, since it was derived from the analysis of the data performed by Emmett: so we are not in fact justified in using the same data to test the hypothesis. The example is chosen, however, merely to exhibit the type of calculation required; and we shall proceed as if the hypothesis had been made *a priori*.

Suppose that $L_{(1)}$, $V_{(1)}$ and $C_{(1)} = L_{(1)}L'_{(1)} + V_{(1)}$ are approximations to L, V and C. We then carry out an iterative procedure in order to obtain better approximations, $L_{(2)}$, $V_{(2)}$ and $C_{(2)}$. We calculate, in turn, the matrices $L'_{(1)}V^{-1}_{(1)}$, $J_{(1)} = L'_{(1)}V^{-1}_{(1)}L_{(1)}$, $(I+J_{(1)})^{-1}$ and $L'_{(1)}V^{-1}_{(1)}A$. Hence we derive $M'_{(1)} = L'_{(1)}C^{-1}_{(1)}A$ in the form $(I+J_{(1)})^{-1}L'_{(1)}V^{-1}_{(1)}A$. An approximation to K is then given by $K_{(1)} = L'_{(1)}C^{-1}_{(1)}(C_{(1)}-A)C^{-1}_{(1)}L_{(1)}$, which is calculated in the form

$$K_{(1)} = (L'_{(1)} - M'_{(1)})(V^{-1}_{(1)}L_{(1)})(I+J_{(1)})^{-1}.$$

The complete matrix $K_{(1)}$ is not of immediate interest; we require, however, an estimate $k^{(1)}_{23}$ of k_{23}, the only non-zero element of K in this case. Let us denote the three rows of $L'_{(1)}$, in order, by $l'_{1(1)}$, $l'_{2(1)}$ and $l'_{3(1)}$, and those of $M'_{(1)}$ by $m'_{1(1)}$, $m'_{2(1)}$ and $m'_{3(1)}$. Then the non-zero elements of $L'_{(2)}$, our next approximation to L', are the corresponding elements of the three row vectors $m'_{1(1)}$, $m'_{2(1)} + k^{(1)}_{23}l'_{3(1)}$ and $m'_{3(1)} + k^{(1)}_{23}l'_{2(1)}$. Finally, the next approximation to V is given by

$$V_{(2)} = \mathrm{diag}(A - L_{(2)}M'_{(1)}).$$

This method appears to converge, and is in principle similar to that of Howe. His method probably converges a little faster. Something like it would be obtained if in the next iteration we found $J_{(2)}$ and $K_{(2)}$ in the same way as before, but defined $M_{(2)}$, differently, by

$$M'_{(2)} = J^{-1}_{(2)}(L'_{(2)}V^{-1}_{(2)}A - M'_{(1)}).$$

It would then, as a matter of fact, be more convenient to calculate $K_{(2)}$ as

$$\{(M'_{(2)} - L'_{(2)})V^{-1}_{(2)}L_{(2)} - K_{(1)}\}J^{-1}_{(2)}.$$

(*c*) *Numerical Example.* A full description of the data to be used is unnecessary. They consist of the correlation coefficients given in Table VII of the paper by Emmett. (For convenience one of the tests has been omitted, and the remaining eight arranged in a different order.) The size of the sample was 211. The coefficients may be regarded as constituting a standardized sample covariance matrix A. This matrix, with lower non-diagonal elements omitted, was as follows:

1·000	0·691	0·679	0·149	0·409	0·382	0·346	0·270
	1·000	0·791	0·285	0·505	0·443	0·471	0·355
		1·000	0·314	0·472	0·372	0·426	0·254
			1·000	0·470	0·385	0·434	0·218
				1·000	0·680	0·639	0·504
					1·000	0·576	0·452
						1·000	0·395
							1·000

The initial trial values for L and V were derived from Emmett's rotated loadings; and, starting from these, a number of iterations was performed. The last of these is reproduced below, where $L'_{(1)}$ and $V_{(1)}$ represent the values obtained from the preceding iteration:

$$L'_{(1)} \quad \begin{matrix} 0{\cdot}500 & 0{\cdot}563 & 0{\cdot}439 & 0{\cdot}358 & 0{\cdot}807 & 0{\cdot}765 & 0{\cdot}663 & 0{\cdot}607 \\ 0{\cdot}623 & 0{\cdot}659 & 0{\cdot}738 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0{\cdot}179 & 0{\cdot}328 & 0{\cdot}558 & 0{\cdot}325 & 0{\cdot}183 & 0{\cdot}349 & 0 \end{matrix}$$

$$V_{(1)} \quad \begin{matrix} 0{\cdot}3619 & 0{\cdot}2177 & 0{\cdot}1570 & 0{\cdot}5605 & 0{\cdot}2431 & 0{\cdot}3813 & 0{\cdot}4386 & 0{\cdot}6316 \end{matrix}$$

$$L'_{(1)}V_{(1)}^{-1} \quad \begin{matrix} 1{\cdot}382 & 2{\cdot}586 & 2{\cdot}796 & 0{\cdot}639 & 3{\cdot}320 & 2{\cdot}006 & 1{\cdot}512 & 0{\cdot}961 \\ 1{\cdot}721 & 3{\cdot}027 & 4{\cdot}701 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0{\cdot}822 & 2{\cdot}089 & 0{\cdot}996 & 1{\cdot}337 & 0{\cdot}480 & 0{\cdot}796 & 0 \end{matrix}$$

$$J_{(1)} = \begin{bmatrix} 9{\cdot}4023 & 4{\cdot}6286 & 3{\cdot}7100 \\ 4{\cdot}6286 & 6{\cdot}5364 & 2{\cdot}0836 \\ 3{\cdot}7100 & 2{\cdot}0836 & 2{\cdot}1879 \end{bmatrix} \qquad (I+J_{(1)})^{-1} = \begin{bmatrix} 0{\cdot}18874 & -0{\cdot}06737 & -0{\cdot}17563 \\ -0{\cdot}06737 & 0{\cdot}18600 & -0{\cdot}04317 \\ -0{\cdot}17563 & -0{\cdot}04317 & 0{\cdot}54629 \end{bmatrix}$$

$$L'_{(1)}V_{(1)}^{-1}A \quad \begin{matrix} 8{\cdot}069 & 9{\cdot}553 & 9{\cdot}182 & 5{\cdot}658 & 9{\cdot}626 & 8{\cdot}529 & 8{\cdot}333 & 6{\cdot}279 \\ 7{\cdot}005 & 7{\cdot}935 & 8{\cdot}264 & 2{\cdot}595 & 4{\cdot}451 & 3{\cdot}747 & 4{\cdot}024 & 2{\cdot}733 \\ 3{\cdot}140 & 4{\cdot}021 & 4{\cdot}201 & 3{\cdot}045 & 4{\cdot}041 & 3{\cdot}372 & 3{\cdot}636 & 2{\cdot}245 \end{matrix}$$

$$M'_{(1)} \quad \begin{matrix} 0{\cdot}4995 & 0{\cdot}5622 & 0{\cdot}4384 & 0{\cdot}3583 & 0{\cdot}8072 & 0{\cdot}7651 & 0{\cdot}6631 & 0{\cdot}6067 \\ 0{\cdot}6238 & 0{\cdot}6587 & 0{\cdot}7372 & -0{\cdot}0300 & 0{\cdot}0049 & -0{\cdot}0232 & 0{\cdot}0301 & -0{\cdot}0116 \\ -0{\cdot}0042 & 0{\cdot}1763 & 0{\cdot}3256 & 0{\cdot}5577 & 0{\cdot}3248 & 0{\cdot}1824 & 0{\cdot}3491 & 0{\cdot}0057 \end{matrix}$$

$$K_{(1)} = 10^{-4}\begin{bmatrix} 0 & 8 & -2 \\ 8 & -11 & 36 \\ -2 & 36 & 4 \end{bmatrix} \qquad k_{23}^{(1)} = 0{\cdot}0036$$

$$L'_{(2)} \quad \begin{matrix} 0{\cdot}4995 & 0{\cdot}5622 & 0{\cdot}4384 & 0{\cdot}3583 & 0{\cdot}8072 & 0{\cdot}7651 & 0{\cdot}6631 & 0{\cdot}6067 \\ 0{\cdot}6238 & 0{\cdot}6593 & 0{\cdot}7384 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0{\cdot}1787 & 0{\cdot}3283 & 0{\cdot}5577 & 0{\cdot}3248 & 0{\cdot}1824 & 0{\cdot}3491 & 0 \end{matrix}$$

$$V_{(2)} \quad \begin{matrix} 0{\cdot}3614 & 0{\cdot}2181 & 0{\cdot}1566 & 0{\cdot}5606 & 0{\cdot}2429 & 0{\cdot}3814 & 0{\cdot}4384 & 0{\cdot}6319 \end{matrix}$$

$V_{(1)}$ and $V_{(2)}$ are diagonal matrices; we have written them as above to save space. The convergence, as in many cases, is extremely slow. It would therefore be unwise to conclude that we are close to the true solution L', in spite of the small difference between $L'_{(1)}$ and $L'_{(2)}$. As Howe has pointed out when discussing the ordinary maximum likelihood equations, it is possible for differences between successive iterates to be extremely small and yet to be far from the exact solution. In practice this does not usually matter, since in such a solution it is unlikely that further iterations would materially improve the fit; and there would thus be little point in trying to obtain an exact solution.

The effect of k_{23} is here not very large. Examination of $K_{(1)}$ shows that the other elements have not yet converged to their final zero values. In the first few iterations the approximations to K were clearly so unreliable that the terms in k_{23} were omitted.

III. Estimation with Oblique Factors

(*a*) *Equations of Estimation.* Now suppose that the factors are oblique, i.e. correlated. Let P be their correlation matrix, of order k. Each diagonal element of P is unity and the typical non-diagonal element is ρ_{rs}, the correlation

coefficient between the rth and sth factors. We may, as before, assume the variance of each factor to be unity.

With the same meaning as before for C, L and V, the log-likelihood function is again given by (1); but C is now expressed in terms of V, L and P by $C = LPL' + V$. We must now maximize (1) not only with respect to the non-zero l_{ir} and to the σ_i^2, but also with respect to the factor correlation coefficients ρ_{rs}. This leads, after some algebra, to three sets of equations. Using circumflex accents to denote estimated values, the first set may be expressed by the statement that those elements of the matrix

$$Q' = \hat{P}(\hat{L}'\hat{C}^{-1} - \hat{L}'\hat{C}^{-1}A\hat{C}^{-1}) \tag{6}$$

which correspond in position to the non-zero elements of L' are to be equated to zero. The second set of equations is

$$\operatorname{diag}(\hat{C}^{-1} - \hat{C}^{-1}A\hat{C}^{-1}) = 0, \tag{7}$$

as in the orthogonal case. The third set may be expressed by the statement that the non-diagonal elements of

$$K = \hat{L}'\hat{C}^{-1}(\hat{C} - A)\hat{C}^{-1}\hat{L} \tag{8}$$

are zero, i.e. that K is a diagonal matrix.

These equations may be simplified in much the same manner as those for the orthogonal case. For convenience we again drop circumflex accents. We first post-multiply (6) by L and obtain $Q'L = PK$. This matrix must have zero diagonal elements, in view of the properties of Q'. But the non-diagonal elements of K are zero, and therefore so also must be its diagonal elements. Thus $K = 0$.

Let us next post-multiply (6) by V, remembering that $V = C - LPL'$ and that $K = 0$. We then obtain

$$Q'V = P(L' - L'C^{-1}A), \tag{9}$$

a matrix which has zeros in the same positions as Q'.

Now post-multiply the matrix on the left-hand side of (7) by V, in the form $C - LPL'$, and use the fact that $\operatorname{diag}(QL') = 0$. The result still has zero diagonal elements, and so $\operatorname{diag}(I - C^{-1}A) = 0$. Pre-multiplication by V then gives

$$\operatorname{diag}(V - A - LPL'C^{-1}A) = 0,$$

which, by use of (9) and of the fact that $\operatorname{diag}(LQ'V) = 0$, is easily seen to be equivalent to

$$V = \operatorname{diag}(A - LPL'). \tag{10}$$

This means that the diagonal elements of C are equal to the corresponding elements of A.

For purposes of computation we may note that the inverse of C is given by

$$C^{-1} = V^{-1} - V^{-1}L(P^{-1} + J)^{-1}L'V^{-1},$$

where $J = L'V^{-1}L$, and that

$$L'C^{-1} = (I + JP)^{-1}L'V^{-1}.$$

Using this result, the equation $K = 0$ may be put in the form

$$P = J^{-1}(L'V^{-1}AV^{-1}L)J^{-1} - J^{-1}. \tag{11}$$

These equations, in different notation, were previously given by Anderson and Rubin and independently by Howe. Our derivation is rather more condensed. Equations (9), (10) and (11) can be solved by an iterative method, though this is even more laborious than that employed in the orthogonal case, since now two matrices of order k require to be inverted in each iteration, J and $I+JP$.

(*b*) *Iterative Method of Solution.* Let $L_{(1)}$, $V_{(1)}$ and $P_{(1)}$ be approximations to L, V and P. Then $C_{(1)} = L_{(1)}P_{(1)}L'_{(1)} + V_{(1)}$ is an approximation to C. We calculate, in turn, the matrices $L'_{(1)}V^{-1}_{(1)}$, $J_{(1)} = L'_{(1)}V^{-1}_{(1)}L_{(1)}$, $J^{-1}_{(1)}$ and

$$(P^{-1}_{(1)} + J_{(1)})^{-1} = P_{(1)}(I + J_{(1)}P_{(1)})^{-1}.$$

After that we find $L'_{(1)}V^{-1}_{(1)}A$, and hence $M'_{(1)} = P_{(1)}L'_{(1)}C^{-1}_{(1)}A$, in the form

$$(P^{-1}_{(1)} + J_{(1)})^{-1}L'_{(1)}V^{-1}_{(1)}A.$$

The elements of $L'_{(2)}$, our next approximation to L', are then obtained by equating those elements of $P_{(1)}L'_{(2)}$ which correspond in position to non-zero elements of L' to the corresponding elements of $M'_{(1)}$.

For $P_{(2)}$, the next approximation to P, we first calculate

$$P^*_{(2)} = (J^{-1}_{(1)}L'_{(1)}V^{-1}_{(1)}AV^{-1}_{(1)}L_{(1)} - I)J^{-1}_{(1)},$$

and then, if necessary, standardize it so that all diagonal elements are unity; this gives $P_{(2)}$. The next approximation to V is given by

$$V_{(2)} = \text{diag}(A - L_{(2)}P_{(2)}L'_{(2)}).$$

As far as the first iteration is concerned this method is identical with that proposed by Howe. For later iterations, however, he suggests a slightly different procedure with the object of increasing to some extent the rate of convergence. If this method were to be used in a second iteration we should carry out the same calculations as before except that we should define $M'_{(2)}$ by

$$M'_{(2)} = J^{-1}_{(2)}(L'_{(2)}V^{-1}_{(2)}A - P^{-1}_{(2)}M'_{(1)}).$$

The method has certain computational advantages, but we have not investigated it very fully.

(*c*) *Numerical Example.* To illustrate the iterative method of solution we use the same data as before. The pattern of factor loadings previously assumed would not now lead to a unique set of loadings, since with oblique factors any non-orthogonal transformation may be applied. Thus the loadings in the first row (or factor) could be replaced by a linear function of those in all three rows, leaving the second and third rows unaltered. To obtain uniqueness we shall make the hypothesis that the pattern of loadings is as follows:

x	x	0	0	x	x	x	x
x	x	x	0	0	0	0	0
0	x	x	x	x	x	x	0

To obtain an initial approximation to L', we pre-multiplied the matrix of final loadings for the orthogonal case by a matrix T given by

$$T = \begin{bmatrix} 1{\cdot}0000 & -0{\cdot}3097 & -0{\cdot}6416 \\ 0 & 1{\cdot}0469 & 0 \\ 0 & 0 & 1{\cdot}1881 \end{bmatrix}.$$

This was chosen so that the third and fourth loadings in the first row are zero, and the diagonal elements of $(TT')^{-1}$ are all unity. The initial approximation to P is given by

$$(TT')^{-1} = \begin{bmatrix} 1{\cdot}000 & 0{\cdot}296 & 0{\cdot}540 \\ 0{\cdot}296 & 1{\cdot}000 & 0{\cdot}160 \\ 0{\cdot}540 & 0{\cdot}160 & 1{\cdot}000 \end{bmatrix},$$

and this enables an approximation for V to be found.

Starting from these values, three iterations were performed. The calculations for the last of these are reproduced below, with $L'_{(1)}$, $V_{(1)}$ and $P_{(1)}$ representing the values obtained from the preceding iteration:

$L'_{(1)}$

0·307	0·245	0	0	0·599	0·648	0·440	0·606
0·653	0·691	0·774	0	0	0	0	0
0	0·211	0·389	0·663	0·385	0·216	0·414	0

$V_{(1)}$

0·3603	0·2171	0·1581	0·5604	0·2430	0·3817	0·4375	0·6328

$L'_{(1)} V_{(1)}^{-1}$

0·852	1·129	0	0	2·465	1·698	1·006	0·958
1·812	3·183	4·896	0	0	0	0	0
0	0·972	2·460	1·183	1·584	0·566	0·946	0

$$J_{(1)} = \begin{bmatrix} 4{\cdot}1375 & 1{\cdot}3362 & 1{\cdot}9702 \\ 1{\cdot}3362 & 7{\cdot}1721 & 2{\cdot}5760 \\ 1{\cdot}9702 & 2{\cdot}5760 & 3{\cdot}0706 \end{bmatrix} \qquad J_{(1)}^{-1} = \begin{bmatrix} 0{\cdot}35047 & 0{\cdot}02215 & -0{\cdot}24345 \\ 0{\cdot}02215 & 0{\cdot}20096 & -0{\cdot}18280 \\ -0{\cdot}24345 & -0{\cdot}18280 & 0{\cdot}63523 \end{bmatrix}$$

$$P_{(1)} = \begin{bmatrix} 1{\cdot}000 & 0{\cdot}297 & 0{\cdot}542 \\ 0{\cdot}297 & 1{\cdot}000 & 0{\cdot}152 \\ 0{\cdot}542 & 0{\cdot}152 & 1{\cdot}000 \end{bmatrix} \qquad (P_{(1)}^{-1} + J_{(1)})^{-1} = \begin{bmatrix} 0{\cdot}18796 & -0{\cdot}00863 & -0{\cdot}04523 \\ -0{\cdot}00863 & 0{\cdot}14802 & -0{\cdot}08314 \\ -0{\cdot}04523 & -0{\cdot}08314 & 0{\cdot}28295 \end{bmatrix}$$

$L'_{(1)} V_{(1)}^{-1} A$

3·896	4·529	3·939	2·906	5·664	5·212	4·764	3·996
7·336	8·308	8·644	2·714	4·659	3·924	4·212	2·863
3·710	4·751	4·962	3·605	4·781	3·989	4·303	2·656

$M'_{(1)}$

0·5012	0·5647	(0·4413)	(0·3597)	0·8082	0·7654	0·6645	0·6062
0·7438	0·7957	0·8330	(0·0769)	(0·2433)	(0·2042)	(0·2246)	(0·1685)
(0·2636)	0·4487	0·5072	0·6630	0·7093	0·5667	0·6519	(0·3327)

$L'_{(2)}$

0·3074	0·2454	0	0	0·6000	0·6489	0·4406	0·6062
0·6525	0·6908	0·7738	0	0	0	0	0
0	0·2107	0·3896	0·6630	0·3841	0·2150	0·4131	0

$$P^*_{(2)} = \begin{bmatrix} 1{\cdot}0009 & 0{\cdot}2978 & 0{\cdot}5423 \\ 0{\cdot}2978 & 0{\cdot}9999 & 0{\cdot}1514 \\ 0{\cdot}5423 & 0{\cdot}1514 & 0{\cdot}9993 \end{bmatrix} \qquad P_{(2)} = \begin{bmatrix} 1{\cdot}0000 & 0{\cdot}2977 & 0{\cdot}5423 \\ 0{\cdot}2977 & 1{\cdot}0000 & 0{\cdot}1515 \\ 0{\cdot}5423 & 0{\cdot}1515 & 1{\cdot}0000 \end{bmatrix}$$

$V_{(2)}$

0·3603	0·2171	0·1581	0·5604	0·2425	0·3814	0·4378	0·6325

As an example of the way in which the non-zero elements of $L'_{(2)}$ are obtained from the elements of $M'_{(1)}$, consider the values in the first column and denote them by x_1 and x_2. Then these values are given by the equations

$$\begin{bmatrix} 1{\cdot}000 & 0{\cdot}297 \\ 0{\cdot}297 & 1{\cdot}000 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0{\cdot}5012 \\ 0{\cdot}7438 \end{bmatrix}.$$

Hence $x_1 = 0{\cdot}3074$ and $x_2 = 0{\cdot}6525$. Similarly, for the values in the fifth column we have

$$\begin{bmatrix} 1{\cdot}000 & 0{\cdot}542 \\ 0{\cdot}542 & 1{\cdot}000 \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0{\cdot}8082 \\ 0{\cdot}7093 \end{bmatrix},$$

and so $x_1 = 0{\cdot}6000$ and $x_3 = 0{\cdot}3841$. In each case the matrix on the left-hand side is either $P_{(1)}$ or a principal minor of $P_{(1)}$. The bracketed values in $M'_{(1)}$ are not required here.

It is difficult to say how near the above values are to the exact solutions of the equations of estimation, but they are sufficiently good for our purposes. It is evident that further iterations would not appreciably improve the fit.

IV. Tests of Hypotheses

We now discuss the problem of testing the hypotheses made in the foregoing sections. Our discussion overlaps that of Howe, but differs in certain details. The general hypothesis with which we are concerned is that the p test variates depend upon k factors, where k is determined in advance, and that certain specified loadings are zero. This is equivalent to a restriction on the form of the covariance matrix C. The alternative hypothesis imposes no such restriction.

Whether the factors are orthogonal or oblique, it is easy to show that the usual approximate χ^2 criterion, based on the likelihood ratio, takes the form

$$n\{\log_e(|\hat{C}|/|A|) + \mathrm{tr}(A\hat{C}^{-1}) - p\}, \tag{12}$$

where we revert to the use of circumflex accents. By making use of the equations of estimation, this may be simplified to

$$n \log_e(|\hat{C}|/|A|). \tag{13}$$

Bartlett [2] has pointed out that, where $k = 0$, the χ^2 test is improved by using, in place of n, the multiplying factor $n' = n - \frac{1}{6}(2p + 5)$. We shall here use it when $k > 0$, though the correct factor is still a matter for conjecture.

At this point it is convenient to discuss orthogonal factors and oblique factors separately. Let us suppose, first, that the factors are orthogonal. If the estimates available are exact or nearly exact solutions of the equations of estimation, the value of χ^2 can be found from (13), with n replaced by n'. The calculation of the determinant of $\hat{C}$ can be considerably simplified if we note that $\hat{C} = \hat{L}\hat{L}' + \hat{V}$, and that the first k latent roots of $\hat{V}^{-1}\hat{C}$ are the same as those of $I + \hat{L}'\hat{V}^{-1}\hat{L} = I + J$, while the remaining latent roots are unity. This means that $|I+J| = |\hat{V}^{-1}\hat{C}| = |\hat{C}|/|\hat{V}|$ and hence that $|\hat{C}| = |\hat{V}|\,|I+J|$.

If, on the other hand, the estimates are not very close to the true solutions of the equations of estimation, using (13) may give a highly misleading value for χ^2. In such cases it is suggested that (12) should be employed instead, with n' in place of n. The calculation of $\mathrm{tr}(A\hat{C}^{-1})$ can be simplified by using the identity $\hat{C}^{-1} = \hat{V}^{-1} - \hat{C}^{-1}\hat{L}\hat{L}'\hat{V}^{-1}$. From this it follows that

$$A\hat{C}^{-1} = (A - M\hat{L}')\hat{V}^{-1},$$

where $M' = \hat{L}'\hat{C}^{-1}A$; and hence that

$$\mathrm{tr}(A\hat{C}^{-1}) = \mathrm{tr}(A\hat{V}^{-1}) - \mathrm{tr}(\hat{L}'\hat{V}^{-1}M).$$

This reduces to p when maximum likelihood estimates are employed. It will be noted that the matrix M' arises in the course of each iteration.

When n is reasonably large an approximation may be found which makes the value of χ^2 very much easier to calculate. Making use of the above result for $A\hat{C}^{-1}$, we have

$$\log_e(|\hat{C}|/|A|) = -\log_e|\hat{C}^{-1}A| = -\log_e|\hat{V}^{-1}(A - \hat{L}M')| = -\log_e|I + X|,$$

where $X = \hat{V}^{-1}(A - \hat{L}M' - \hat{V})$. The diagonal elements of X are zero. The non-diagonal elements, x_{ij}, are (apart from the factor $\hat{V}^{-1}$) similar to the residuals usually calculated; for large n they will tend to be small. Hence the above expression is approximately equal to $\frac{1}{2}\mathrm{tr}(X^2) = \sum_{i<j}(x_{ij}x_{ji})$. Thus, when n is large, we may use this quantity multiplied by n' as an approximation to χ^2.

If m is the number of non-zero loadings and if these are uniquely determined, the number of degrees of freedom for χ^2 is $\frac{1}{2}p(p-1) - m$. This is because there are $\frac{1}{2}p(p+1)$ variances and covariances, and there are p residual variances and m loadings to be estimated.

In the case of oblique factors the same results still apply, apart from certain simple alterations. Since $\hat{C} = \hat{L}\hat{P}\hat{L}' + \hat{V}$, we now have $|\hat{C}| = |\hat{V}|\,|I + J\hat{P}|$. The expressions for $A\hat{C}^{-1}$ and $\mathrm{tr}(A\hat{C}^{-1})$ are exactly the same as before if we now define M' to be $\hat{P}\hat{L}'\hat{C}^{-1}A$, which is how it appears in the iterations. The approximate expression for χ^2 is also calculated in the same way as before. The number of degrees of freedom for χ^2 is, however, reduced by $\frac{1}{2}k(k-1)$, since we have to estimate the factor correlation coefficients. It now becomes $\frac{1}{2}(p-k)(p+k-1) - m$.

Values of χ^2 were calculated for the data previously analysed, though a χ^2 test is not really justifiable here. For these data we have $n = 210$, $p = 8$, $k = 3$, and we take our multiplying factor n' to be $n - \frac{1}{6}(2p+5) = 206{\cdot}5$. In the orthogonal case the number m of non-zero loadings is 17, so the number of degrees of freedom is 11. The value of χ^2 was found by the approximate method to be 5·0, which, taken at its face value, corresponds to a probability of 0·93. This high probability is not surprising since, as already remarked, the hypothesis set up was derived from a previous analysis of the data. In the oblique case the value of χ^2 was almost the same as before, but now $m = 15$ and the number of degrees of freedom is 10.

V. Summary and Conclusions

In this paper we have demonstrated that, for various hypotheses, it is possible to solve numerically the maximum likelihood equations of estimation. But it is evident that the labour involved would increase enormously as the number of factors increased. In fact, if $k > 3$, an electronic digital computer would become almost a necessity if a large number of iterations were to be performed. As convergence usually appears to be rather slow, a large number *would* have to be performed to obtain an exact solution. On the other hand, it seems likely that in practice two or three iterations would be sufficient to obtain a reasonably good fit and to give a valid test of significance.

It would be possible to outline methods for finding large-sample standard errors of the various estimated values and of the residual covariances. This has not been done since such methods would, unfortunately, be of no practical use owing to their extreme complexity. One possibility for the future would be the development of methods of estimation which, though not fully efficient in the statistical sense, would only involve calculations of a relatively simple kind. It is fairly easy to invent such methods, but far more difficult to assess their efficiency; and, until this is done, they would not have much value from a statistical standpoint.

References

[1] Anderson, T. W., and Rubin, H. (1955). Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, V, iii.

[2] Bartlett, M. S. (1950). Tests of significance in factor analysis. *Brit. J. Psychol. Stat. Sect.*, III, 77.

[3] Emmett, W. G. (1949). Evidence of a space factor at 11+ and earlier. *Brit. J. Psychol. Stat. Sect.*, II, 3.

[4] Howe, W. G. (1955). Some contributions to factor analysis. *Report No.* ORNL-1919, Oak Ridge National Laboratory (Ph.D. thesis, University of North Carolina).

[5] Lawley, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proc. roy. Soc. Edin.*, LX, 64.

[6] Thurstone, L. L. (1947). *Multiple-Factor Analysis.* Chicago: University of Chicago Press.

Vol. XI Part I The British Journal of Statistical Psychology May 1958

THE RELATION BETWEEN COEFFICIENTS OF CORRELATION AND DIFFICULTY FACTORS

By Harvey F. Dingman
University of Southern California

Summary. The purpose of the following inquiry was to test the view, put forward by Ferguson and others, that unequal dichotomization of items tends to produce spurious difficulty factors. Subtests of three levels of difficulty were compiled for each of the three factors, and applied to nearly 500 students. The correlations were calculated by the phi, cosine–pi, and full tetrachoric formulae, and four factors extracted. The results obtained do not support Ferguson's hypothesis.

I. Problem

Ferguson's work [7] has suggested that, when phi is used as a coefficient of fourfold correlation and the frequency distributions of the several items are unequally dichotomized, spurious factors may appear on factorizing the correlation table in the usual way[1]. As a rule the spurious factors are related not to any concrete psychological trait, but to a manifest bias in the size of the phi. Accordingly, Burt [2, 3], Wherry and Gaylord [17], Gourlay [8], and others have recommended the substitution of the tetrachoric coefficient (r_t). Since r_t does not exhibit the systematic bias of the phi, it should not (so it is argued) give rise to any spurious factor. Carroll [6], however, holds that the use of r_t may still produce such a factor unless the data are first corrected for guessing.

Guilford [9] published, about the same time as Ferguson [7], an analysis revealing factors associated with the difficulty of different items in Seashore's test of musical talent. The correlations (r_t) were estimated from the graphs of Chesire, Saffir, and Thurstone [5]. It might therefore seem that, when computed in this way, r_t is suspect. In the introduction to their tables, Thurstone and his colleagues contend that this procedure does not yield estimates that are systematically biased. For general purposes Burt calculated the values for r_t from Pearson's tables [15]; but, when the lines of division ran close to the median, he relied on Sheppard's simpler cosine–pi formula. The bias in estimates calculated by the latter procedure has recently been tabled by Bouvier *et al.* [1].

The purpose of the present investigation is to examine the extent of what are commonly described as 'spurious difficulty factors', and to estimate their influence on other factors in the same analysis. To do so, it is necessary to select

[1] According to Gourlay the idea of a 'difficulty factor' seems to have first been mentioned in Burt's earlier publications ([3], [8], and refs.). They were found both when correlating persons and when correlating tests: e.g. in factorizing item-correlations for the Binet tests, they appeared to indicate the relative difficulty of items assigned to different ages.

items of varying difficulty and to construct relatively univocal tests for the various factors. If factors emerge corresponding with the level of difficulty found for the items, we may assume that a so-called difficulty factor is present. If, however, the loadings correspond with psychological factors independently demonstrated, we may conclude that 'real' factors have emerged without appreciable bias and without suppression or distortion by any element of spuriousness in the coefficients of correlation.

II. Method

Fairly pure tests for each of three factors—Spatial Visualization, Verbal Comprehension, and Logical Evaluation—were selected from those constructed for Guilford's studies of high level aptitudes and fully described in a series of publications [10, 11, 12, 13, 14], to which the reader is referred for fuller details. The level of difficulty of the individual items was determined from these previous researches. Suitable items were then chosen with a view to securing three subtests for each factor. Each subtest was specially constructed so as to be relatively homogeneous in item difficulty; and the items themselves were selected so that the average level of difficulty, as judged by the percentages passing and failing, should be 0·90, 0·70, and 0·50 respectively. There were thus nine tests in all—three for every factor—each of a different level of difficulty.

The tests were administered to 479 male college students. Table I gives the parameters obtained for the several tests. A Pearson tetrachoric coefficient was calculated for every possible intercorrelation (9×9 in all). To check Carroll's hypothesis, the scores were also corrected for guessing. The items all involved multiple choice with three alternatives; and the standard formula was used. The corrected scores were then correlated by the ordinary Pearson product-moment procedure.

Table I. Observed and Hypothetical Test Parameters

Factor Tested	Average Item Difficulty	Test Mean Without Correction for Guessing	Test Mean With Correction for Guessing	Point of Dichotomy Without Correction for Guessing	Point of Dichotomy With Correction for Guessing
1. Spatial Visualization	0·90	7·24	6·54	0·91	0·89
4. Verbal Comprehension	0·90	9·42	9·18	0·93	0·88
7. Logical Reasoning	0·90	7·67	6·90	0·89	0·91
2. Spatial Visualization	0·70	6·05	5·04	0·67	0·71
5. Verbal Conprehension	0·70	7·82	6.86	0·74	0·74
8. Logical Reasoning	0·70	5·38	3·86	0·61	0·73
3. Spatial Visualization	0·50	3·48	2·20	0·56	0·48
6. Verbal Comprehension	0·50	6·48	3·14	0·52	0·51
9. Logical Reasoning	0·50	4·09	1·78	0·56	0·52

Note.—For the cosine–pi coefficient all the dichotomies were at $0{\cdot}50 \pm 0{\cdot}02$ except that for test 4. This was dichotomized at 0·68 owing to its skewed distribution.

The distributions of the scores, both those corrected for guessing and those not corrected, were dichotomized at the point of average difficulty, as determined from the previous sample. Phi coefficients were computed, and the r_t coefficients estimated from Thurstone's graphs. Correlations were also calculated for score distributions dichotomized at the medians; here the cosine–pi approximation was adopted. There were thus eight correlation matrices to be analysed, one for each of the four correlation coefficients under each of the two conditions, with correction for guessing and without (cf. Table II).

The correlation matrices were factorized separately by the complete Thurstone centroid process [16], iterations being continued until the communalities become stable. Since three 'real' common factors were present in each matrix, with the possibility of a spurious factor in addition, four factors in all were extracted from each matrix. The four factors were then rotated orthogonally by the Zimmerman graphic process [18]. Improvement in the three 'real' or 'psychological' factors was alone considered when making the rotations, since there was no desire to 'improve' or to destroy the fourth or 'spurious' factor.

III. Results

Ferguson's main contention is that, 'since tests used in many batteries differ substantially in difficulty for the population tested, the factors resulting from the differences in the nature of the tasks will be confused with factors resulting from differences in their difficulty'. From the construction of the present batteries we should expect (i) three factors to result from 'the *nature* of the tasks' and possibly (ii) one factor to result from 'the differences in their *difficulty*'. Hence there are two questions to be decided: (i) are the psychological factors (or what are taken to be such) likely to be difficulty factors in disguise? (ii) Is a difficulty factor likely to be mistaken for a psychological factor?

Table II. Factor Loadings for the Psychological Factors (Means for Eight Analyses)

	r	r_t	r_c	ϕ
Without Correction for Guessing	0·666	0·741	0·699	0·527
With Correction for Guessing	0·644	0·712	0·696	0·513

Note.—r = product moment coefficient; r_t = tetrachoric coefficient; r_c = cosine–pi coefficient; ϕ = phi-coefficient.

(i) In the present case three apparently 'real' factors were found in each of the eight matrices. There is an occasional lack of orthogonality, e.g. in the case of the harder test for Verbal Comprehension and the easier test for Logical Reasoning. But in each analysis the several factors are readily identifiable psychologically, in the manner anticipated. It is unnecessary to tabulate here the detailed loadings for every test and every factor. The average figures are

given in Table II. The hypothetical loadings we should expect from the differences in item difficulty (first column of Table I) will be found in the first column of Table III. On comparing the size of the hypothetical figures with those of the averages for the 'real' or psychological factors, it seems clear that there is little or no likelihood of the differences in difficulty producing a 'spurious factor' large enough to be confused with a genuine psychological factor. Hence, in the matrices here factorized, if any factor due to the irregular dichotomies or the differences in item difficulty were present, we should expect it to appear as a separate fourth factor.

Table III. Hypothetical and Actual Loadings for Difficulty Factor

Test	Hypothetical Loadings	Actual Loadings: Scores uncorrected for Guessing				Scores corrected for Guessing			
		r	r_t	ϕ	r_c	r	r_t	ϕ	r_c
1	0·20	0·14	0·21	0·08	0·22	0·08	−0·08	0·13	0·16
2	0·20	0·13	0·38	0·25	0·25	0·28	0·03	0·22	0·27
3	0·20	0·32	0·03	0·15	−0·14	0·30	0·41	0·19	−0·14
4	0·00	−0·05	0·20	−0·17	0·18	0·06	0·00	−0·22	0·16
5	0·00	0·02	−0·20	−0·05	−0·16	0·03	0·15	0·01	−0·15
6	0·00	−0·08	−0·21	0·13	−0·19	0·05	0·24	−0·04	−0·17
7	−0·20	−0·23	−0·21	0·26	−0·25	−0·16	−0·09	−0·23	−0·15
8	−0·20	−0·11	−0·18	−0·21	−0·09	0·13	0·18	−0·22	−0·12
9	−0·20	−0·10	0·05	−0·12	0·15	0·00	−0·19	+0·07	+0·16
Product Sum		0·0206	0·0192	0·0110	0·0104	0·0138	0·0092	0·0184	0·0132

Note.—The hypothetical loadings (i.e. the loadings expected on Ferguson's hypotheses, shown in the first column) were selected so as to be in keeping with the figures in Table I.

(ii) In Table III the first column shows the hypothetical loadings we should expect for a difficulty factor on the basis of Ferguson's arguments. These are derived from the figures for average difficulty (first column of Table I) on the assumption that the factor emerging will be of the ordinary bipolar type. The remaining columns give the actual loadings. They are all comparatively small. Judging by mere inspection, the first set (r, uncorrected for guessing) shows some agreement both in signs and in size. Ferguson, however, used the phi coefficient without correction for guessing. And here certainly the pattern of loadings (column 4) does not conform with what we should anticipate on his hypothesis. The last line in the table gives the product-sums: for r, r_t, and ϕ (all without correction) these imply correlations of 0·90, 0·63, and 0·37; for r_t (with correction) only 0·32.

IV. Conclusion

Results derived from a single set of tests are admittedly not decisive; but it is hoped that the present experiment may stimulate other investigators to study the problem along the lines suggested. Meanwhile we may legitimately conclude

that, in exploratory analyses at any rate, there is little need for anxiety about the possibility of confusing difficulty factors with psychological factors. In certain investigations (as in Burt and John's studies of the relative difficulty of the Binet tests, cited by Gourlay) it may even be desirable to consider difficulty as a separate factor. In that case the product-moment coefficient without correction for guessing would seem to be the most effective. When the difficulty factor is likely to be spurious, r_t with correction for guessing would appear to provide the most prudent procedure.

References

[1] Bouvier, E. A., Perry, N. C., Michael, W. B., and Hertzka, A. F. (1954). A study of the error in the cosine–pi approximation to the tetrachoric coefficient of correlation. *Educ. Psychol. Measmt.*, XIV, 715–721.

[2] Burt, C. (1921). *Mental and Scholastic Tests.* London: Staples.

[3] Burt, C. (1940). *The Factors of the Mind.* University of London Press.

[4] Burt, C. (1951). A reply to Gourlay. *Brit. J. Psychol. Stat. Sect.*, IV, 65, 76.

[5] Chesire, L., Saffir, M., and Thurstone, L. L. (1933). *Computing Diagrams for the Tetrachoric Correlation Coefficient.* Chicago: University of Chicago Bookstore.

[6] Carroll, J. B. (1945). The effect of difficulty and chance success on correlations between items or between tests. *Psychometrika*, X, 1–20.

[7] Ferguson, G. A. (1941). The factorial interpretation of test difficulty. *Psychometrika*, VI, 323–329.

[8] Gourlay, N. (1951). Difficulty factors arising from the use of tetrachoric correlations in factor analysis. *Brit. J. Psychol. Stat. Sect.*, II, 65–76.

[9] Guilford, J. P. (1941). The difficulty of a test and its factor composition. *Psychometrika*, VI, 67–77.

[10] Guilford, J. P., Comrey, A. L., Green, R. F., and Christensen, P. R. (1950). A factor-analytic study of reasoning abilities, I. Hypotheses and descriptions of tests. *Rep. Psychol. Lab.*, No. 1. Los Angeles: University of Southern California.

[11] Guilford, J. P., Wilson, R. C., Christensen, P. R., and Lewis, D. J. (1951). A factor-analytic study of creative thinking, I. Hypotheses and description of tests. *Rep. Psychol. Lab.*, No. 4. Los Angeles: University of Southern California.

[12] Guilford, J. P., Green, R. F., Christensen, P. R., Hertzka, A. F., and Kettner, N. W. (1952). A factor-analytic study of Navy reasoning tests with the Air Force Aircrew Classification Battery. *Rep. Psychol. Lab.*, No. 6. Los Angeles: University of Southern California.

[13] Guilford, J. P., Hertzka, A. F., Berger, R. M., and Christensen, P. R. (1952). A factor-analytic study of evaluative abilities, I. Hypotheses and description of tests. *Rep. Psychol. Lab.*, No. 7. Los Angeles: University of Southern California.

[14] Guilford, J. P., Berger, R. M., and Christensen, P. R. (1954). A factor-analytic study of planning, I. Hypotheses and description of tests. *Rep. Psychol. Lab.*, No. 10. Los Angeles: University of Southern California.

[15] Pearson, K. (1932). *Tables for Statisticians and Biometricians*, Pt. II. Cambridge University Press.

[16] Thurstone, L. L. (1947). *Multiple Factor Analysis.* Chicago: University of Chicago Press.

[17] Wherry, R. J., and Gaylord, R. H. (1944). Factor pattern of test items and tests as a function of the correlation coefficient: content, difficulty, and constant error factors. *Psychometrika*, IX, 237–244.

[18] Zimmerman, W. S. (1946). A simple graphical method for orthogonal rotation of axes. *Psychometrika*, XI, 51–55.

Vol. XI Part I The British Journal of Statistical Psychology May 1958

SECOND ORDER FACTORS AND THE ORGANIZATION OF COGNITIVE FUNCTIONS

By G. Bernyer
Centre National de la Recherche Scientifique, Paris

Summary. The object of the following inquiry is to compare the results obtained by factorizing the same set of data according to (i) Thurstone's method of ' oblique simple structure ' and (ii) Burt's method of ' overlapping ' and ' subdivided ' group factors. For this purpose the correlations between Thurstone's oblique factors were made the basis of an analysis into what he has called ' second order factors '; and a simple working procedure was devised which enables the investigator to transform the oblique factors into a factor pattern including both general and group factors.

Two batteries of tests, differing slightly in their composition, have been applied to two different groups. With both sets of data much the same factors were ultimately obtained with both procedures; and the factor pattern was fully consistent with Burt's theory of a ' hierarchical classification ' leading to ' subdivided group factors '. Thus the results obtained by Thurstone's procedure, when carried to its logical completion, so far from being incompatible with Burt's theory of a hierarchical classification by subdivided factors, tends to confirm it[1].

I. The Alternative Methods

There are many different methods of factorial analysis, but nearly all begin in much the same way: to account for a set of observed correlations between a number of variables in terms of a minimum number of common factors, they commonly start by an analysis into what Burt has called ' general and bipolar ' factors, as in his method of ' simple summation ' and Thurstone's ' centroid method '. The later stages, however, may diverge in consequence of the different procedures used in rotating these initial factors. The choice of the rotation depends essentially on the principle of classification adopted *a priori* by the investigator; and the final solutions may be divided into three main categories:

(A) A classification of the different variables into separate clusters corresponding to independent factors—the so-called ' simple orthogonal structure ' preferred by many American authors (e.g. Guilford).

(B) A classification on two levels, comprising (i) a general factor and (ii) a set of group factors which may partly overlap (Burt's method of group factor analysis).

[1] The substance of the following paper is based on an investigation reported more fully, with diagrams and other details, in *Biotypologie* [2]. The author is deeply indebted to the authorities of the *Centre National de la Recherche Scientifique* and to the editor of that paper for permission to publish this account and to Dr. Charlotte Banks for assistance with the translation.

(C) A hierarchical classification comprising (i) a general factor as before, with (ii) two or more broad group factors which are then subdivided into (iii) narrower group factors, these again often being further subdivided in their turn (Burt's method of subdivided factors).

Thurstone's method of analysis into oblique factors has often been considered, quite wrongly, as an analysis of type (A), since most authors using it limit their inquiry to first order factors, and consequently their analyses end with a classification of the variables into clusters. However, by introducing correlations between the factors, the method admits, at least by implication, a definite organization among these factors: that is to say, it really involves the presence of latent factors which are much broader than the group factors explicitly extracted, and are in effect superposed on them. The method is thus really akin to possible analyses of type (B) or (C), and it would therefore seem more logical to regard it merely as a preliminary stage which permits the investigator to proceed to a solution of one or other of these types. Actually, however, extremely few investigations of this kind have been undertaken, no doubt for reasons which are partly theoretical and partly practical. To begin with, Thurstone's disciples deny that any psychological interest can attach to a general factor, and consider his unitary or primary abilities as the only valid kind; consequently they rarely push their analysis as far as factors of the second order, or, if they do, they are merely content to extract oblique factors of the second order, so as to avoid the introduction of any general factor. But secondly, those investigators who might feel tempted to extract a general factor of the second order seem often to have encountered obstacles which they treat as insurmountable. It is this second difficulty I wish more particularly to examine here.

However, one or two recent researches have revealed the close agreement between the factorial results obtained by the methods of Burt and of Thurstone respectively. Reuchlin [8], for example, has re-analysed the data obtained by Moursy [6], which were specially planned to verify the existence of a hierarchical structure, and he shows that it is quite possible to recover such a structure by carrying the analysis to a second or third stage. In the present study I shall describe two cases in which Thurstone's method of analysis has been carried to factors of the second order, and exactly the same data factorized by the methods proposed by Burt. In doing so I shall endeavour to deduce certain conclusions regarding the relative advantages and applicability of the two procedures.

II. Factors of the First Order

The full details of the experimental and statistical procedure have been described in two earlier papers [1, 2]. The subjects tested were two groups of adolescents aged about 15 years—pupils in a school for apprentices, numbering 160 (group A) and 156 (group B). The batteries used in the two cases were slightly different. As will be seen from the tables (I and III), the original scheme included 23 tests; but of these only 16 were used with the first group, and 15 with the second, 11 tests being identical in both cases. For detailed descriptions of

the several tests, and the tables of observed correlations, the reader may refer to the first of the two papers just cited (1).

The results of the centroid analysis are reproduced in Tables I and III, and the structures based on the reference vectors (obtained by rotation to oblique axes) in Tables II and IV. The loadings (or ' saturations ') which are larger than 0·30 are printed in bold type. With the first battery five oblique factors were obtained, and of these, four re-appeared in the analysis of the second battery. The following is a brief description of their nature.

(i) Factor *R* appears to be a factor of *general and verbal reasoning*. Since it enters into two composite tests of intelligence, it may be considered as itself a complex factor comprising such components as deduction, induction, and verbal comprehension. It is no doubt related to what French has called a ' factor of schooling ' and to the verbal-scholastic factor (*v* : *ed*) of certain English factorists.

(ii) Factor S_2 corresponds with the factor which Thurstone describes as *ability in visualizing a configuration in which there is a movement or displacement of certain parts*. Reference to the tests suggests that it is partly dependent on the mechanical experience acquired by the subjects, and may therefore in some degree be related to the mechanical factor, *m*.

(iii) Factor S_1 corresponds with what Thurstone describes as *skill in visualizing a rigid structure when moved to different positions*. It is in fact the familiar ' space factor ', often designated (after El Koussy) by the symbol *K*.

(iv) Factor *A* may be identified with what Burt, Harrell, and others have termed a factor for *manual dexterity*.

(v) Factor *F* may be defined as *skill in resolving problems by concrete manipulations through trial and error*. It resembles what Alexander calls the factor of ' practical intelligence '.

III. Analyses of the First Battery (Group A)

1. *Group Factor Analysis*. The analysis into general and group factors was carried out by the method described by Burt [5]. According to the principles he lays down, the sign-pattern displayed by the bipolar factors provides an objective indication of how the tests are to be classified. The signs suggested an initial division into four main groups; and this formed the basis of the first trial. The results so obtained, however, suggested the need for adding a fifth group. The formulae for the direct procedure yield non-overlapping factors. We begin by making a direct analysis into non-overlapping factors (superposed on the general or basic factor). As a rule, however, without some overlap one seldom obtains a complete fit for all the observed correlations. This preliminary scheme is therefore taken as the basis for calculating an orthogonal rotation matrix which is applied direct to bipolar centroid factors and transforms them into a set of overlapping group factors by a rigorous arithmetical procedure. A few supplementary adjustments were applied to minimize the non-significant

TABLE I. FACTOR SATURATIONS FOR GROUP A

Tests	Bipolar Factors I	II	III	IV	V	Group Factors g	r	s_2	s_1	a	f
1. Logical intelligence	0·53	0·29	−0·40	0·06	−0·12	**0·40**	**0·59**	0·17	−0·06	0·05	−0·03
2. Intelligence test (W)	0·46	0·34	−0·42	−0·04	−0·16	**0·36**	**0·62**	0·09	0·00	−0·02	−0·09
6. Progressive matrices	0·43	0·28	−0·34	−0·21	0·10	**0·37**	**0·51**	−0·12	0·06	−0·04	0·12
15. Technical intelligence	0·68	0·29	0·09	0·26	−0·11	**0·61**	0·20	**0·47**	−0·07	0·00	0·06
13. Mechanical movement (IT_1)	0·59	0·42	0·21	0·26	−0·10	**0·59**	0·12	**0·49**	−0·07	−0·18	0·06
16. Mechanical ability	0·64	0·09	0·19	0·23	−0·15	**0·56**	0·03	**0·45**	0·02	0·13	0·02
19. Assembly: blocks	0·48	0·05	0·27	0·21	0·33	**0·45**	−0·13	0·22	−0·11	0·07	**0·42**
18. Wiggly blocks	0·49	0·08	0·22	−0·04	0·23	**0·49**	−0·05	0·08	0·07	0·03	**0·30**
11. Form relations	0·59	0·12	0·07	−0·37	0·14	**0·60**	0·14	−0·12	0·30	0·02	**0·20**
9. Assembly: surfaces	0·65	0·17	0·15	−0·31	−0·16	**0·67**	0·13	0·08	**0·36**	−0·02	−0·04
12. Designs from memory	0·49	−0·08	0·15	−0·16	−0·28	**0·44**	0·01	0·15	**0·31**	0·18	−0·17
17. Minnesota formboard	0·60	−0·27	0·14	−0·35	0·16	**0·53**	−0·06	−0·15	**0·35**	**0·34**	0·22
20. Winding test	0·37	−0·27	0·05	0·18	−0·14	0·22	−0·04	0·25	0·02	**0·39**	−0·05
21. Piorkowski rings	0·47	−0·31	0·21	0·18	0·11	**0·35**	−0·19	0·21	0·02	**0·41**	0·20
23. Manual skill	0·39	−0·56	−0·29	0·14	0·24	0·11	0·12	−0·06	−0·09	**0·72**	0·25
22. Walther discs	0·34	−0·66	−0·20	0·08	−0·06	0·07	0·02	0·01	0·06	**0·77**	−0·02
Variance per cent	27·3	10·0	5·7	4·7	3·2	21·2	7·3	5·9	3·0	10·2	3·3

TABLE II. FACTOR SATURATIONS FOR GROUP A

Tests	Simple Structure R	S_2	S_1	A	F	Second Order Group Factors g	r	s_2	s_1	a	f	Subdivided Factors g	r	s_2	p	s_1	a	f
1. Logical intelligence	**0·58**	0·18	−0·03	0·05	−0·03	**0·45**	**0·57**	0·14	−0·03	0·01	−0·06	**0·46**	**0·56**	0·13	0·01	−0·03	0·01	−0·06
2. Intelligence test (W)	**0·57**	0·07	0·03	−0·05	−0·08	**0·44**	**0·57**	0·05	−0·01	−0·08	−0·10	**0·46**	**0·56**	0·04	−0·01	0·01	−0·07	−0·11
6. Progressive matrices	**0·51**	−0·06	0·20	−0·08	0·18	**0·37**	**0·49**	−0·06	0·12	−0·09	0·15	**0·37**	**0·49**	−0·07	0·04	0·12	−0·09	0·15
15. Technical intelligence	0·21	**0·60**	−0·04	0·09	0·06	**0·55**	0·21	**0·54**	−0·02	0·02	0·05	**0·58**	0·20	**0·51**	0·00	−0·03	0·06	0·05
13. Mechanical movement (IT_1)	0·11	**0·62**	−0·03	−0·09	0·07	**0·52**	0·12	**0·57**	−0·03	−0·15	0·06	**0·57**	0·10	**0·54**	−0·09	−0·02	−0·08	0·06
16. Mechanical ability	0·05	**0·55**	0·03	0·21	0·02	**0·51**	0·04	**0·49**	0·05	0·16	0·02	**0·51**	0·04	**0·48**	0·11	0·02	0·18	0·01
19. Assembly: blocks	0·01	**0·49**	0·02	0·16	**0·45**	0·24	−0·02	**0·44**	0·05	0·11	**0·44**	0·24	−0·01	**0·43**	0·06	0·02	0·13	**0·44**
18. Wiggly blocks	0·03	0·28	0·23	0·07	**0·37**	**0·34**	0·00	0·26	0·19	0·05	**0·36**	**0·34**	0·01	0·26	0·11	0·16	0·06	**0·36**
11. Form relations	0·15	0·02	**0·52**	0·00	**0·31**	**0·53**	0·13	0·01	**0·37**	−0·01	0·30	**0·50**	0·14	0·02	0·21	**0·34**	−0·01	0·30
9. Assembly: surfaces	0·06	0·13	**0·51**	−0·02	0·05	**0·68**	0·06	0·12	**0·35**	−0·04	0·05	**0·66**	0·06	0·11	0·20	**0·33**	−0·02	0·05
12. Designs from memory	−0·07	0·13	**0·33**	0·18	−0·13	**0·50**	−0·06	0·10	0·26	0·17	−0·13	**0·46**	−0·05	0·11	0·25	0·22	0·16	−0·12
17. Minnesota formboard	0·00	−0·01	**0·54**	**0·32**	**0·33**	**0·44**	−0·03	−0·03	**0·44**	**0·33**	**0·32**	**0·36**	0·01	0·00	**0·40**	**0·35**	0·27	**0·33**
20. Winding test	−0·01	0·27	−0·04	**0·42**	−0·08	0·22	−0·02	0·22	0·03	**0·40**	−0·07	0·17	0·01	0·23	0·21	−0·03	0·36	−0·08
21. Piorkowski rings	−0·07	**0·37**	0·05	**0·47**	0·21	0·23	−0·10	**0·31**	0·11	**0·44**	0·21	0·18	−0·07	**0·32**	0·26	0·03	**0·40**	0·20
23. Manual skill	**0·31**	0·06	−0·07	**0·72**	0·25	0·02	0·25	−0·01	0·05	**0·71**	0·23	−0·07	0·30	0·03	**0·36**	−0·06	**0·59**	0·22
22. Walther discs	0·11	0·02	−0·01	**0·75**	−0·04	0·08	0·07	−0·05	0·10	**0·76**	−0·04	−0·03	0·13	−0·01	**0·43**	−0·01	**0·64**	−0·6
Variance per cent.	—	—	—	—	—	17·7	6·5	8·2	3·8	10·3	4·3	17·1	6·6	7·7	4·8	2·8	7·6	4·3

TABLE III. FACTOR SATURATIONS FOR GROUP B

	Tests	Bipolar Factors I	II	III	IV	Group Factors g	r	s_2	a
2.	Intelligence test (W)	0·54	0·33	−0·37	−0·20	**0·42**	**0·62**	0·04	−0·09
3.	Burt's reasoning test	0·48	0·30	−0·38	−0·13	**0·34**	**0·60**	0·08	−0·06
4.	Superfluous words	0·41	0·25	−0·21	−0·09	**0·32**	**0·40**	0·10	−0·05
6.	Progressive matrices	0·46	−0·07	−0·13	−0·17	**0·39**	0·26	−0·09	0·18
13.	Mechanical movement (IT_1)	0·69	0·36	0·16	0·16	**0·61**	0·19	**0·49**	0·00
14.	Mechanical movement (IT_2)	0·59	0·35	0·29	0·09	**0·61**	0·04	**0·44**	−0·08
16.	Mechanical ability	0·64	0·23	0·22	0·18	**0·57**	0·08	**0·44**	0·09
10.	Arranging cubes	0·70	0·12	0·20	0·06	**0·65**	0·10	**0·31**	0·15
19.	Assembly: objects	0·58	0·07	0·17	−0·07	**0·58**	0·08	0·15	0·11
18.	Wiggly blocks	0·55	−0·20	0·24	−0·04	**0·54**	−0·08	0·05	**0·32**
9.	Assembly: surfaces	0·58	−0·05	0·16	−0·15	**0·60**	0·05	0·02	0·17
17.	Minnesota formboard	0·70	−0·34	0·02	−0·07	**0·58**	0·13	−0·05	**0·50**
21.	Piorkowski rings	0·33	−0·14	0·07	−0·13	**0·33**	0·02	−0·08	0·17
23.	Manual skill	0·30	−0·58	−0·27	0·28	−0·03	0·15	−0·03	**0·74**
22.	Walther discs	0·17	−0·63	−0·07	0·33	−0·08	−0·09	0·00	**0·72**
	Variance per cent	28·8	10·0	4·9	2·6	23·2	7·3	5·2	10·6

TABLE IV. FACTOR SATURATIONS FOR GROUP B

	Tests	Simple Structure R	S_2	S_1	A	Second Order Group Factors g	r	s_2	s_1	a	Subdivided Factors g	r	s_2	$p+f$	s_1	a
2.	Intelligence test (W)	**0·60**	−0·01	0·03	−0·03	**0·49**	**0·57**	−0·02	−0·01	−0·09	**0·51**	**0·55**	−0·01	0·00	0·02	−0·02
3.	Burt's reasoning test	**0·57**	0·02	−0·04	0·03	**0·44**	**0·54**	−0·01	−0·07	−0·06	**0·46**	**0·53**	0·01	−0·01	−0·02	0·02
4.	Superfluous words	**0·37**	0·07	0·00	0·00	**0·39**	**0·35**	0·04	−0·02	−0·05	**0·40**	**0·34**	0·05	0·01	0·00	0·00
6.	Progressive matrices	0·29	−0·08	0·24	0·03	**0·34**	0·27	−0·05	0·21	0·17	**0·31**	0·26	−0·06	0·27	0·15	0·03
13.	Mechanical movement (IT_1)	0·06	**0·51**	−0·03	0·05	**0·73**	0·08	**0·34**	−0·01	−0·01	**0·70**	0·07	**0·37**	0·13	−0·02	0·03
14.	Mechanical movement (IT_2)	−0·07	**0·48**	0·05	−0·08	**0·67**	−0·04	**0·33**	0·08	−0·10	**0·66**	−0·06	**0·35**	0·09	0·04	−0·07
16.	Mechanical ability	−0·04	**0·48**	0·02	0·08	**0·65**	−0·02	**0·32**	0·04	0·07	**0·61**	−0·02	**0·35**	0·20	0·02	0·05
10.	Arranging cubes	0·02	**0·36**	0·16	0·04	**0·67**	0·04	0·24	0·18	0·13	**0·62**	0·03	0·26	0·30	0·11	0·03
19.	Assembly: objects	0·05	0·19	0·24	−0·05	**0·54**	0·06	0·14	0·24	0·08	**0·51**	0·04	0·14	0·27	0·16	−0·04
18.	Wiggly blocks	−0·08	0·12	**0·34**	0·05	**0·43**	−0·06	0·10	**0·34**	0·29	**0·36**	−0·08	0·09	**0·45**	0·22	0·04
9.	Assembly: surfaces	0·06	0·08	**0·35**	−0·07	**0·50**	0·07	0·07	**0·34**	0·14	**0·46**	0·05	0·06	**0·35**	0·23	−0·05
17.	Minnesota formboard	0·15	0·00	**0·38**	0·19	**0·48**	0·15	0·01	**0·36**	**0·48**	**0·38**	0·14	0·00	**0·60**	0·25	0·16
21.	Piorkowski rings	0·05	−0·05	0·27	−0·02	0·24	0·06	−0·02	0·26	0·15	0·21	0·04	−0·03	0·27	0·18	−0·02
23.	Manual skill	0·14	−0·03	−0·01	**0·62**	0·03	0·12	−0·06	−0·03	**0·75**	−0·12	0·16	−0·03	**0·54**	−0·01	**0·50**
22.	Walther discs	−0·10	0·03	0·01	**0·57**	−0·06	−0·10	−0·01	0·01	**0·72**	−0·21	−0·07	0·01	**0·53**	0·01	**0·45**
	Variance per cent	—	—	—	—	23·7	5·8	2·8	3·9	10·1	21·7	5·6	3·2	10·8	1·7	3·3

saturations; and the final factor pattern thus obtained is given in Table I (cols. 6–11).

As the table shows, the general or ' basic ' factor (g) has positive saturations for all the tests, fully significant in every case except two and sometimes rising to well over 0·60. As a result, it accounts for 21·2 per cent of the total variance, that is, for about 40 per cent of the common factor variance. The five group factors are associated with much the same clusters of tests as the five factors in the oblique analysis, and may be interpreted in an analogous fashion. We have used the same symbols to describe them, substituting small letters for capitals in order to distinguish the orthogonal group factors from the oblique. With the present battery, and judged by the contributions to the total variance, the factor a (manual dexterity) is the largest of the group factors, and factor r (verbal reasoning) and s_2 (perception of mechanical differences) the next largest: s_1 is the smallest.

2. *Analysis into Second Order Group Factors.* Thurstone introduced the further stages into his analysis in the hope of reconciling his own point of view with Spearman's. This would have been effected readily enough had the analysis led to a *single* general factor of the second order. In practice, however, most of those who have sought to extract such a factor have encountered two main difficulties. First, the application of Spearman's formula often leads to impossible results, for instance, to a so called 'Heywood case' in which the squares of the saturations are negative or larger than one. Secondly, in the vast majority of instances a single factor by itself proves insufficient to account for the inter-correlations between the first order factors, and often leaves unexplained residuals. Nevertheless, both difficulties can usually be resolved; and, by using the procedure adopted here, it is almost always possible to transform the second order factors into a factor pattern comprising both general and group factors.

First of all, instead of estimating the general factor saturations by Spearman's formula, which is apt to yield decidedly unsatisfactorily figures when only a small number of variables is concerned, I have preferred to determine for the factor g those values which have the greatest likelihood, and for this purpose adopted the working procedure described by Lawley [7]. The results are given in Table V (cols. 1–6). In tables of this type (V–VIII) the correlations between the first-order factors (first 4 or 5 columns) are shown below the empty diagonal, the residual correlations above the diagonal, and the factor saturations in the last one or two columns of this set; and the complete analysis, i.e. the saturations of the first order factors with the second order factors, are shown in the set of columns on the right.

In Table V, it will be seen, several of the residuals are by no means negligible. They can be accounted for by assuming a slight overlap between the specifics. For this purpose the table of residual correlations has been itself analysed into five centroid factors; and these have been rotated so as to yield a minimum of overlapping. The complete analysis of the five first-order factors is shown in Table V (cols. 7–12).

Here, it will be observed, the factor *g* is not a perfectly general factor, but shows a slight bipolarity. In the saturations for the tests (Table II, col. 5) the bipolarity disappears, owing no doubt to the fact that those which have saturations for factor *F* also have saturations for either S_1 or S_2. The analysis of first order factors, however, is of interest because it reveals a distinct antithesis between abstract and concrete intelligence respectively. This in turn reflects the fact

TABLE V. GROUP A. FIRST ANALYSIS OF FIRST ORDER FACTORS

	R	S_2	S_1	*A*	*F*	*g*	*g*	*r*	s_2	s_1	*a*	*f*
R	—	−0·01	−0·01	−0·08	−0·09	**0·49**	**0·49**	**0·87**	−0·03	−0·03	−0·04	−0·04
S_2	**0·32**	—	0·00	−0·13	0·00	**0·67**	**0·67**	0·02	**0·74**	0·00	−0·08	0·00
S_1	**0·40**	**0·57**	—	0·07	0·02	**0·84**	**0·84**	0·00	0·00	**0·54**	0·00	0·00
A	−0·08	−0·14	0·07	—	0·01	−0·01	−0·01	−0·04	−0·08	0·14	**0·99**	0·00
F	−0·30	−0·29	**−0·34**	0·02	—	**−0·42**	**−0·42**	−0·06	0·00	0·04	0·00	**0·90**

TABLE VI. GROUP A. FINAL ANALYSIS OF FIRST ORDER FACTORS

	R	S_2	S_1	*A*	*F*	I	II	*g*	*r*	s_2	*p*	s_1	*a*	*f*
R	—	−0·03	0·00	−0·01	0·09	**0·48**	−0·12	**0·50**	**0·87**	0·04	0·00	−0·03	−0·04	−0·05
S_2	**0·32**	—	0·01	−0·01	0·01	**0·68**	−0·21	**0·72**	−0·01	**0·70**	−0·03	0·00	0·00	0·00
S_1	**0·40**	**0·57**	—	0·00	0·01	**0·86**	0·12	**0·80**	0·02	0·01	**0·33**	**0·50**	0·01	0·01
A	−0·08	−0·14	0·07	—	−0·02	−0·01	**0·57**	−0·15	0·03	−0·02	**0·55**	−0·01	**0·82**	−0·01
F	−0·30	−0·29	**−0·34**	0·02	—	**−0·42**	0·06	**−0·42**	−0·05	0·02	−0·05	0·01	−0·01	**0·91**

TABLE VII. GROUP B. FIRST ANALYSIS OF FIRST ORDER FACTORS

	R	S_2	S_1	*A*	*g*	*g*	*r*	s_2	s_1	*a*
R	—	0·01	−0·06	−0·11	**0·64**	**0·64**	**0·76**	−0·03	−0·05	−0·10
S_2	**0·59**	—	0·02	−0·07	**0·90**	**0·90**	0·03	**0·43**	0·04	−0·05
S_1	**0·35**	**0·59**	—	**0·43**	**0·63**	**0·63**	0·02	0·04	**0·64**	**0·44**
A	−0·14	−0·12	**0·40**	—	−0·05	−0·05	−0·02	−0·05	−0·01	**1·00**

TABLE VIII. GROUP B. FINAL ANALYSIS OF FIRST ORDER FACTORS

	R	S_2	S_1	*A*	I	II	*g*	*r*	s_2	*p*+*f*	s_1	*a*
R	—	0·01	−0·01	0·02	**0·54**	**−0·38**	**0·66**	**0·75**	0·00	0·00	0·00	0·00
S_2	**0·59**	—	0·01	−0·01	**0·80**	**−0·38**	**0·87**	0·01	**0·47**	0·15	0·01	−0·01
S_1	**0·35**	**0·59**	—	0·01	**0·86**	0·28	**0·54**	−0·02	0·01	**0·73**	**0·42**	0·01
A	−0·14	−0·12	**0·40**	—	0·22	**0·73**	−0·25	0·03	−0·01	**0·73**	0·00	**0·64**

that with the present sample of pupils—a somewhat selected group consisting of young adolescents who have just left the primary school—those that are least gifted for intellectual and scholastic tasks are apparently best endowed with *F* (' manipulative intelligence ').

In Table V a slight overlapping will also be remarked between factors s_1 and *a*. From this table, by using Thurstone's procedure, we can readily derive a

transformation matrix which enables us to transform the ' simple structure ' based on five reference vectors into a set of test-saturations for orthogonal factors of the second order.

The analysis of the tests into second order group factors is shown in Table II. The results are here closely analogous to those obtained by the more direct method of Burt. For each group factor we find the same cluster of tests in each of the two cases. The only difference discernible is that, in the second analysis, tests belonging to group s_2 and some of the tests belonging to group f have rather larger saturations with the group factors and lower saturations for g. Thus, although the method of extracting factors has apparently some slight influence on the detailed figures, the two procedures have essentially the same outcome.

3. *Analysis into Subdivided Factors of the Second Order.* The data available in the present research do not permit of a direct analysis into subdivided factors in accordance with Burt's method [3]. Nevertheless, it is tempting to inquire whether a more complete analysis of the correlations between the first order factors would not furnish somewhat fuller information about the organization of the factors than is supplied by the preceding analysis, in which, owing to the limitations of the data, we have been obliged to adopt a somewhat arbitrary factorial model. Accordingly, starting from the correlations between the five first-order factors (Table VI, cols. 1–4, below the diagonal), two second-order factors (I, II) have been extracted by Lawley's method[1]. They seem to account quite satisfactorily for the intercorrelations between these five factors. Table VI shows (cols. 1–4, above the diagonal) the residuals left after extracting these two factors, and in the last two columns the factor saturations.

A diagram plotting the saturations for the second-order factors is instructive (see [2], Figs. 1 and 2); and by its aid it is possible to carry out a graphic rotation by the method commonly used by Thurstone. With this procedure the choice of axes is based, not on the factor pattern, but on psychological considerations. As we have seen, the factor R represents a broad factor closely related to Burt's ' intellectual ' factor. It may therefore be plausibly inferred that g would suffice to account for the correlations of R with the other factors. This suggests passing the axis for the factor g through R, and taking for the second factor an axis p, perpendicular to g. Although the residuals are rather small, we have also introduced, as in the preceding analysis, a slight overlapping between the specifics to secure a second-order analysis statistically equivalent to the centroid analysis. The complete analysis of the first-order factors is shown in Table VI.

[1] It will be remembered that the ' simple summation ' formula put forward by Burt (termed by Thurstone the ' centroid formula ') was a simplified substitute for the more elaborate method of ' weighted summation ', which fits the successive factors according to the principle of least squares. In the method of least squares, however, when the observed values differ in ' precision ', they should first be weighted accordingly. This method of ' correcting ' the final values was suggested by Burt ([4], pp. 113f.) who proposed a rather laborious method of successive approximation. The formulae themselves are algebraically equivalent to those reached by adopting the method of maximum likelihood; and Lawley's working method seems the simplest for practical purposes.

The new factor g is near enough to the old factor g, the general or ' basic ' factor obtained in the group factor analysis; and its bipolarity in Table VI (col. 8) can be explained in much the same way as before.

The results thus obtained indicate the need for introducing certain minor adjustments into the factor saturations shown for the second order group factors in Table II so as to allow the introduction of a factor for p. This leads to the modified factor pattern shown in the last six columns of table II. Here the broad factor p does not exhibit quite the same breadth as one would expect for what Burt has called the factor of ' practical ability '. But this no doubt is due largely to the fact that our tests are not numerous enough to yield a complete scheme of subdivided factors. However, p here overlaps two narrower group factors, the spatial factor s_1 and the factor for manual dexterity, a; and, on referring to the tests for which it has appreciable saturations, we also note that quickness of perception seems to play some part. Possibly, therefore, the factor might be identified with the *perceptual factor* found in several factorial analyses.

IV. Analyses of the Second Battery (Group B)

The correlations obtained from the second group of examinees have been analysed in the same way as the preceding. It will therefore be sufficient briefly to summarize the main results.

1. *Group Factor Analysis.* A preliminary analysis into non-overlapping group factors was first carried out by dividing the tests into four groups in accordance with the bipolar sign pattern. With one group, the residuals left after the general factor had been extracted were not fully significant; and the analysis was therefore repeated with only three group factors. The final results, with a general factor and overlapping group factors, are set out in Table III.

Here the general factor accounts for 23·2 per cent of the total variance: that is, for about 50 per cent of the common factor variance. Factors a, r, and s_2 are as conspicuous as before; r and s_2 have saturations for the same tests as the oblique factors R and S_2 in Table IV. But the spatial factor s_1, which was previously the smallest, has apparently disappeared. The tests belonging to the cluster for that factor are now strongly saturated with factors g and a. This probably results from the inappropriate composition of the tests included in this cluster, which were perhaps unduly diversified.

2. *Analysis into Second Order Group Factors.* Starting with the matrix of the correlations between the four oblique factors, we determined as before the most likely values for the common factor g. Table VII (cols. 1–5) shows the correlations between the factors of the first order, the residual correlation after g has been extracted, and the saturations of the four factors with g.

It is plain that one second order factor alone can not suffice to explain the correlations between the first-order factors. There remains, for instance, a large residual correlation between S_1 and A which is unaccounted for. Nevertheless, if we insist on keeping to the factorial model chosen in advance, it is

unallowable to introduce a second common factor, and the only alternative is to assume that the specifics themselves are correlated.

Proceeding as in the case of the first battery, we have analysed the four oblique factors into orthogonal factors of the second order as shown in Table VII (cols. 6–10). Here the residual correlation between s_1 and a is accounted for by postulating an overlap between these two factors. From this analysis of the first-order factors it is possible to deduce the saturations of the tests for the group factors of the second order as shown in Table IV.

The general factor, g, has approximately the same size as in the factor analysis carried out by Burt's procedure; but it differs somewhat in its nature. Here the saturations for g are higher for the tests belonging to the groups S_2 and R, and lower for the tests belonging to group S_1, and a group-factor for S_1 can now be obtained. Consequently the variances for the different group factors differ somewhat in the two analyses.

3. *Analysis into Subdivided Factors of the Second Order.* As before, we have gone on to analyse the correlations between the four first-order factors and have found two common factors. Table VIII (cols. 1–6) shows the intercorrelations, residuals, and factor-saturations. Once again we have undertaken a graphical rotation of the second-order factors, by passing the axis for the first factor through R. As with the first battery, we obtain a group factor which includes both S_1 and A. The resulting analysis of the oblique factors is shown in Table VIII (cols. 7–12). and the analysis into subdivided factors in the last 5 columns of Table IV.

On comparing the figures with those obtained for the first battery, we observe that f (practical intelligence) no longer appears as a separate factor, but is now merged with p (the perceptual factor), with which it largely overlapped in the former case: p (now fused with f), however, still appears as a broad group factor, subdivided into the two narrower factors s_1 and a.

V. Conclusions

1. First of all, it will be seen, the factors ultimately obtained with the two different methods of factor analysis—Thurstone's and Burt's—are very much the same. Thus, even if the investigator begins by obtaining a ' simple structure ' of oblique factors, it is nearly always possible, by extending the calculations to an analysis of the second order and applying a simple working procedure, to transform the oblique structure into a factor pattern containing both a general factor and group factors. Thus interpreted, Thurstone's oblique structure, so far from being incompatible with the classification of abilities proposed by Burt, actually tends to confirm his theory of a hierarchical structure for the factors of the mind.

2. Secondly, the factors obtained from the two different groups, even with batteries differing somewhat in the number and nature of the tests employed, appear to be very much the same. Thus the general and group factors extracted seem to be invariant.

3. Burt's method is more direct. But the adoption of the more circuitous procedure is of special interest when it is desirable to demonstrate a more complex organization of factors than Thurstone's theory of a simple structure in itself suggests. On the other hand, the direct analysis into subdivided factors by Burt's method requires a rather varied assortment of tests, in order that divergent lines of classification can be displayed at the outset even in the bipolar analysis; and consequently, if the battery to be analysed has not been suitably planned or selected, his method may prove inapplicable. In such cases, when certain factors are missing from the battery, and when one no longer obtains a hierarchical subdivision in its classical form, on analysis into oblique factors may still frequently bring to light this type of organization among the fundamental factors.

References

[1] Bernyer, G. (1956). Analyses factorielles de deux batteries de tests appliquées aux candidats d'une Ecole de Mécanique, *Biotypologie*, XVII, 177–204.

[2] Bernyer, G. (1957). Organisation des fonctions cognitives et facteurs de second ordre. *Biotypologie*, XVIII, 21–39.

[3] Burt, C. (1949). Subdivided factors. *Brit. J. Psychol. Stat. Sect.*, II, 41–63.

[4] Burt, C. (1949). Alternative methods of factor analysis. *Brit. J. Psychol. Stat. Sect.* II, 98–121.

[5] Burt, C. (1950). Group factor analysis. *Brit. J. Psychol. Stat. Sect.* III, 40–75.

[6] Moursy, E. M. (1952). The hierarchical organization of cognitive levels. *Brit. J. Psychol. Stat. Sect.*, V, 151–180.

[7] Lawley, D. N. (1942). The application of the maximum likelihood method to factor analysis. *Brit. J. Psychol.*, XXXIII, 172–175.

[8] Reuchlin, M. (1955). Facteurs observés et facteurs théoriques en psychologie. Ap. *L'analyse factorielle et ses applications* 187–205. Paris : Centre National de la Recherche Scientifique.

JOURNAL COMMITTEE

A meeting of the Journal Committee was held on 25 October, 1957. The Editor outlined the proposals in regard to contents and layout of future issues, and the Treasurer of the Society reported the decisions of the Council regarding the finances of the journal and the proposal to publish three parts per annum instead of two. It was stated that the changes in printer and style of printing, recommended in the Editor's memorandum of 1954, had at length been sanctioned by Council; and the result was that, for the first time since the journal started, the heavy deficits had now been converted into profits on both parts of the recent volume. As announced in the previous number, Council had agreed that such profits should be used to issue a third part together with an index to the first ten volumes. The Treasurer, however, had since found himself faced with a much larger expenditure for the activities of the Society as a whole; and Council had now decided that the profits must be placed to the credit of the Society's general account.

The Committee expressed its regret at the decision. It was explained that this, as well as the earlier decision of Council to discontinue the journal if the deficit continued, had greatly diminished the number of contributions now coming in. Accordingly the Committee instructed the Editor to ask the President to clarify certain points in respect of the Society's future policy; and recommended that, if a supplementary part proved impracticable, the regular issues should be appreciably enlarged, in order that subscribers should at least receive the same amount of material as before. With the same purpose in view, it was also proposed that the section for ' Notes and Correspondence ' (in small type) should be revived; and that, following these changes, fresh publicity leaflets should be prepared stating that the journal had now achieved a stable position.

The Committee approved the provisional list of contents for the forthcoming numbers. The Editor explained that, in accordance with the Committee's previous recommendations, it was planned to devote the first issue largely to the subject of ' methodology ', and much of the second issue to the ' theory of games ' and ' linear programming '. The material already set up for the supplementary part of the preceding volume would now be included in the forthcoming number.

To increase circulation in this country it was agreed that the wisest policy would be to lay chief stress on the ' expository and educative function ' of the journal, as proposed when it was first established, i.e. to publish articles, discussions, and shorter notes on topics of current interest to subscribers, to draw attention to new or neglected developments in statistical method, and to ventilate and answer criticisms, rather than to aim solely at papers of an original or highly technical character. Since this imposed a heavy burden on the Editors, it was suggested that, as with the British Journal of Educational Psychology, profits might be used to pay specialists to contribute material for which subscribers had asked. Articles requiring elaborate tables and formulae should for the present be deferred, since they were expensive to print, and often unintelligible to the majority of British subscribers.

The Committee asked that these proposals should be brought to the notice of readers and contributors, and that apologies should be conveyed to subscribers and authors for the delays caused by recent changes. The Editor wishes to tender his thanks to those who have already submitted practical suggestions, and hopes that readers will continue to send their views and their criticisms. Our readers will also wish so express their gratitude to a private benefactor whose generosity has made it possible to increase the size of this and subsequent issues.

CHARLOTTE BANKS,
Hon. Secretary.

Vol. XI Part I The British Journal of Statistical Psychology May 1958

DEFINITION AND SCIENTIFIC METHOD IN PSYCHOLOGY

By Cyril Burt
University College, London

La filosofia è scritto in questo grandissimo libro che continuamente ci sta aperto innanzi a gli occhi—io dico, l'universo; ma non si può intendere se prima non s' impara a intender la lingua, e conoscer i caratteri, ne' quali è scritto: egli è scritto in lingua matematica.—Galileo Galilei, *Il Saggiatore* (1623), *Quaest.* 6.

Summary. The recent controversies about definitions in psychology seem to overlook the special nature of definition in empirical science. This is but one of the many instances of the way modern psychology tends to ignore the formal requirements of scientific methodology.

In order to determine the logical structure of an empirical science, a brief review is undertaken of the logical structure of a typical example, namely, mechanics. It appears that the structure consists essentially in a hierarchy of propositions based on a hierarchy of abstract concepts. A distinction is drawn between formal or constitutive definitions and rules of empirical interpretation (including so-called 'operational definitions') and between postulates which do not call for proof and hypotheses which do. The general conditions for valid proof in an inductive science are then outlined, and the criteria for distinguishing a science from a pseudo-science briefly indicated.

The conclusions reached suggest that the recommendations put forward, on *a priori* grounds, by several psychologists who have recently proposed different methodologies of their own are inadequate and misleading. And on the basis of an actual study of working procedures an attempt is made to deduce the chief logical requirements appropriate for a complex science like psychology.

I. The Problem of Scientific Method[1]

The Nature of Definition in Science

In a recent paper [32] Mr. Miles has discussed the interpretations of the term 'intelligence' put forward in this journal and elsewhere; and has made it the text for an instructive enquiry into the problem of definition in psychology. "I hate definitions," says the hero of one of Disraeli's novels; "I hold to the maxim of Thomas à Kempis: 'melius sentire aliquid quam scire eius definitionem'." In much the same spirit the modern psychologist is apt to leave his basic ideas to be grasped through enlightened insight rather than through formal exposition; and in this, as in other respects, the modes of argument

[1] In reviewing the policy of the journal at a recent meeting, the Journal Committee strongly urged that more attention be given to problems of methodology; and, largely no doubt as a result of the discussions initiated by contributors to earlier numbers and to the Society's *Bulletin*, several letters have been received by the editors on issues of this type. The following paper has accordingly been written partly to answer the questions so raised, but chiefly in the hope of provoking a more thorough examination of the whole problem from different points of view. My object, therefore, has been not so much to lay down any final conclusions as to draw fire from opposite camps.

he adopts, particularly when the subject is of popular interest, are, as Mr. Miles observes, conspicuously lacking in the logical rigour found in all other sciences[1].

Mr. Miles' examination of the question is, he says, based largely on the recent work of Robinson [42], and seems admirable so far as it goes. Yet it scarcely goes far enough. Nowhere, I think, does he take into account the special type of definition needed in the empirical sciences, such as nowadays rely mainly on mathematical or statistical techniques.

The Need for a Conceptual System

Mr. Miles begins his criticisms by regretting that "Burt in his article omits to explain what he understands by 'the strict logical sense' of the word definition". He himself recognizes two main types—first definitions that give the meanings of *words* and secondly definitions that describe the *things* for which the words stand. The view that I wish to defend is that definitions are concerned not so much with words or things as with *concepts*. Ideas are born before they are named: and in any branch of empirical science the fundamental problem is to determine what concepts are most appropriate[2].

Now the merits of a particular concept can be assessed only when it is examined in the light of the whole scientific doctrine in which it plays an essential part. And my main objection to the definitions of 'intelligence' and the like, put forward by critics such as Dr. Blackburn and Dr. Heim, is that the writers apparently assume that scientific terms can be defined, so to speak, *in vacuo*, without any reference to the wider conceptual context in which they are to function. Similarly, my chief criticism of Mr. Miles's discussion is that he says little or nothing to rectify this misleading assumption[3].

[1] Several writers have suggested that this unscientific character springs largely from the fact that, in this country at any rate, psychology until recently has been treated as a subject for an Arts rather than for a Science degree; and even today, it is said, "at least half the students receive no training whatever in scientific method". It may, however, be doubted whether even in science the courses are so arranged as to teach methods as distinct from results. In any case, as we shall see, many psychologists hold that psychology itself is not amenable to the methods of the other sciences.

[2] Cf. Sigwart [46], I, p. 286. "A definition is a judgement that states the meaning of a term denoting a concept". The title of my article [13] to which Mr. Miles refers ('Evidence for the *Concept* of Intelligence') and the heading of the opening section ('The History of the *Concept*') were meant to indicate that my chief aim was to formulate an adequate concept.

By the 'logical sense' I meant the sense given to the word in standard textbooks of logic, such as those I quoted, as distinct from the loose popular usage which treats the word 'define' as a synonym for 'describe', 'discuss', or 'explain'.

[3] For Mr. Miles the "problem that is really central . . . is the problem how to establish assertions of the form 'there really *exists* an 'x' when 'x' is not a straightforward concrete noun'". His answer is that "mental factors have a real existence when they are factors of the body", i.e. when they are identifiable with material structures that can be, in principle, seen or touched. This is a doubting Thomas's criterion. It would be accepted by the plain man. But would it be accepted by the ordinary scientist? Do physicists seek to show that 'forces', 'rays of light', 'fields', or the 'waves' of quantum mechanics really 'exist'? And, if not, how exactly do they suppose their concepts *are* related to what is commonly regarded as the world of existence? These are questions which, I think, can only be answered by actually examining the methodology of the various sciences; and only then can we consider whether there is any sense in discussing the 'existence' of 'factors'.

The reason why so many psychologists adopt this attitude is, I imagine, that, unlike other branches of science, modern psychology appears to have left them with no orderly scheme of concepts to fall back on. Indeed, I am tempted to echo the verdict with which William James closed his own *Briefer Course*. "A string of raw facts; a wrangle about opinions, a little classification and generalization on the mere descriptive level; but not a single law in the sense in which physics shows us laws, and scarcely a single proposition from which any practical consequence can be plausibly deduced. This is no science; it is scarcely the hope of a science" (p. 468). And several recent commentators, particularly philosophers or logicians who have paid some attention to the logic of the older sciences, have advanced somewhat similar criticisms[1].

Modern psychology has undergone a number of striking developments since James first penned this stern indictment. Yet each of the newer theories has been condemned in similar terms. Psychoanalysis, Dr. Eysenck tells us, is no authentic science, but merely a counterfeit. Factor analysis, according to Mr. Richmond, falls into precisely the same category. Ward is still more drastic in his rejection of "the whole doctrine of individual or differential psychology and its claims to be a scientific discipline". But how are we to discriminate between what is spurious and what is sound? What criteria are we to apply when we wish to decide whether psychoanalysis, for instance, is just a pseudo-science like astrology, or, as McDougall once averred, "a genuine science at the Galilean stage"? This, in my view, is the crucial problem; and it is this that I want to examine in the pages that follow.

On one preliminary point, I fancy, we should all agree. "What distinguishes science from philosophy, common sense, and sheer superstition", says Poincaré, "is its specific logical procedure—in one word, its methods". And it is the faulty methods, or the lack of method, that has formed the target for attack with each of the critics I have quoted. What then is the methodology which the psychologist should adopt, if his subject is to make good its claim to be a genuine branch of science?

Alternative Proposals

Several writers[2] during recent years have endeavoured to lay down what they regard as the canons of procedure that psychologists should observe, and

[1] Professor Broad, for instance, tells us that "psychology at present hardly deserves the name of a 'science'; and it will remain in this unsatisfactory state until someone succeeds in doing for it what Galileo, Descartes, and Newton did for physics" ([7], pp. 440). Similarly Professor Braithwaite declares that the great defect of psychology as a scientific system is that "it has produced a vast number of hypotheses, but up to the present has hardly attempted to unify them by the use of colligating concepts" ([4], p. 30). Professor Eysenck complains that, in many of its fields, particularly the more complex, psychology "still keeps painfully close to common sense, and has failed signally to elaborate concepts and laws of its own . . . It has charted for itself a dangerous course, threatened by the Scylla of nebulous Freudian generalization, and the Charybdis of pointlessly amassing unrelated 'facts'" ([20], pp. 9–15). And in almost every college common room one may hear the physiologist upbraiding the psychiatrist, and the chemists and physicists rebuking the psychologist, because "neither of them seems to possess the barest inkling of what constitutes scientific proof".

[2] As might perhaps be expected, the earliest instances are to be found in the attempts of mathematical psychologists, chiefly factorists, to adapt the postulational procedure revived by

thus in some measure to remove the weaknesses which the majority acknowledge. Five distinct trends seem discernible.

1. The first is the effort, noted by Mr. Miles, to define every key-word in as objective a manner as possible. This seems to have been largely prompted by Bridgman's demand for a re-examination of current definitions in the physical sciences [5]; and most psychologists who have taken up the problem favour the adoption of his proposals in their own science. All concepts which cannot be 'operationally defined' ('mind', 'consciousness', 'innate intelligence', and most mental 'faculties' or 'capacities') are to be ruthlessly banned; and the more ardent reformers have finished up with a programme of 'physical monism' almost as radical as that of the earlier Viennese positivists[1].

2. A number of psychological textbooks go deeper, and start with the announcement that "the essential postulate of all scientific enquiry is the discovery of causes". The most systematic exposition of this view is that undertaken by Professor F. V. Smith [47]. His motto might well be Virgil's: "Felix qui potuit rerum cognoscere causas". He begins by maintaining that, contrary to the teaching of Mach and Pearson, the aim of the scientist should be "to explain, not merely to describe": and explanation, he contends, consists in "ascertaining and proving" causal relations. His view is supported by a detailed analysis of half a dozen instances of well-established scientific 'explanations', drawn mainly, though not exclusively, from mechanics; and from these he seeks to derive the "logical criteria for satisfactory explanation"[2]. He then applies this touchstone to estimate the success achieved by the six best known psychological 'systems' of the present century. He concludes that each "accepts the position that human behaviour is determined", and that in each "the causal determinants of behaviour are spatially and temporally continuous with it".

With both behaviourists and psychoanalysts the determinism adopted is of the thoroughgoing mechanistic type resolutely proclaimed by almost every physicist of the Victorian age. The psychoanalyst, it is true, is willing to ascribe a causal influence to purely psychical entities, e.g. to the 'unconscious' and its more or less mystical contents. But the behaviourist is resolved to exorcize every vestige of 'the ghost in the machine', and insists that the causal determinants of human behaviour must reside exclusively in the material

mathematicians like Hilbert; in the experimental field the most notable example is to be seen in the work of Clark Hull, particularly his outline of 'A Scientific Theoretical System in Miniature' (*Psych. Rev.*, 1937, XLIV, 1–32).

[1] Cf. S. S. Stevens, 'The operational basis of psychology', *Amer. J. Psychol.*, (1935), XLVII, pp. 323–330, and the symposium with Bridgman and others, *ibid.*, (1945), LII, no. v. For 'physical monism' see Pratt [40], pp. 21f.

[2] They may be summarized as follows. "Description of an observed sequence does not of itself afford the same conviction as causal explanation implying necessity": to demonstrate causation (i) we must establish (*a*) some kind of qualitative and quantitative identity between states succeeding each other in *time* and (*b*) some kind of continuity in *space*; furthermore (ii) the relations so established must be relations not merely of 'regularity' (as Hume and his modern followers suppose) but of 'necessity'.

structures of the human body. For him, as for the earlier logical positivists, 'die physikalische Sprache is die Universalsprache der Wissenschaft'[1].

In what follows we shall have to enquire whether a more systematic analysis of the physical sciences really bears out Professor Smith's interpretation. However, to my mind the fundamental issue is somewhat different from that which he discusses: it is, not so much whether the modern psychological 'systems' accept the same set of concepts and postulates as the older sciences, but whether they can themselves be regarded as scientific 'systems' in the same technical sense as the rest: and this means asking how they set about 'ascertaining and proving' their causal relations or whatever kind of generalization they prefer to substitute.

3. When we take up this question, we find an apparent conflict between those who hold (with Bacon and Mill) that the inductive sciences should begin with *observations* and those who maintain (with Whewell and Jevons) that they should begin with *hypotheses*. Professor Zangwill adheres to the former and older view. For him the first step is "the collection of the facts"; next comes "the devising of hypotheses to explain the facts"; last of all the hypotheses proposed are to be submitted to the test of fact. In short, take care of the facts and the hypotheses will take care of themselves. It is these methodological considerations, he says, that have led him to "dismiss the whole theory and practice of mental testing as mere technology", and factor analysis as "a brilliant but misguided departure from the central path of empirical psychology" ([55], p. 160). And he would, I imagine, commend the injunction offered by Voltaire to the *savants* of his day: "Assurons-nous bien du fait avant de nous inquiéter de la cause"; and the equally pertinent remark which Shaw puts into the mouth of Newton: "I never theorize before I observe; else my observations would be coloured by the lenses through which I see them"[2].

4. On the whole, however, it is the second view that seems of late to have gained the warmest support. "What is needed in psychology", writes Professor Eysenck, "is a more intensive use of the hypothetico-deductive method, in which a clear hypothesis is first stated, deductions are then made, and experiments performed to verify or disprove them" ([20], p. 16). As Dr. Taylor and others have pointed out, the invention of this procedure is commonly ascribed (wrongly, so I hold) to Galileo, and its introduction from physics into psychology is usually credited to Clarke Hull [24]. However, many of the writers who use the phrase do so in a loose and non-technical fashion to cover any type of method that relies primarily on deduction as contrasted with induction, in the old-fashioned senses of those terms—what Professor Popper calls 'deductivism' and 'inductivism' respectively [38]. Whether either of these descriptions offers an accurate account of the procedure actually adopted by modern science is a question we shall have to investigate.

[1] Cf. O. Neurath, 'Einheitswissenschaft und Psychologie', *Einheitswissenschaft*, I, 23-36.
[2] Cf. Newton, *Opticks*, qn. 31 and his four 'Rules of Philosophizing', *Principia*, bk. iii, *ad init.*

5. Finally, there is an active group of writers who protest that "the official programme of psychology—a programme that treats the ordinary behaviour of the ordinary human being as the subject of an esoteric science, with laboratory apparatus and mathematical techniques—is just a misguided endeavour to transport the procedure of the physicist into regions where those methods are entirely out of place" [45]. "The scientific methods that we owe to Galileo and his successors lead", we are told, "to a true description of physical reality, not because of the nature of the methods, but because of the nature of physical reality. To suppose that human personalities can be studied in this way is to commit a glaring blunder. A human personality is something to be 'understood' by the intuitive insight of common sense, not something to be 'explained' by scientific analysis." In short, what is called psychology is "not a science, but an art like medicine or a narrative like history". On this view, of course, the question of scientific methodology does not arise[1].

The Need for an Empirical Study of Scientific Method

The only way to check these alternative accounts is to analyse in logical terms the formal structure of a typical science. Nearly all the writers I have quoted begin by laying down *a priori* what they hold to be the nature of scientific method, and then go on to consider how far this or that science conforms with their preconceptions. I suggest that the opposite approach will be more helpful, namely, to examine those sciences that have proved most successful, and observe at first hand the methods they have worked out[2].

[1] Cf. [22] and [33]. Professor Eysenck rightly observes that "the popularity of this view among non-scientists derives from its non-scientific nature", just as "attacks on mathematical procedures appeal strongly to non-mathematicians". In this country its main champions are, as he says, the clinical psychologists, partly no doubt because they have to report their findings to those who are neither mathematicians nor scientists, but still more, I fancy, because their task is practical rather than theoretical. Nevertheless, the doctrine is not without support from statistical and philosophical writers who cannot be accused of scientific ignorance. Dr. Chambers, for instance, maintains that "it is highly doubtful whether the concept of measurable quantity can be applied at all to mental qualities"; and Mr. Richmond joins him in poking fun at the 'psychometrists' who "treat love as a commodity to be weighed and made up into parcels, tied with pink ribbon" [17, 41]. Professor Ryle argues the same point still more clearly. "The abandonment of the dream of a science of psychology, the counterpart of Newtonian science, must involve the abandonment of any notion that 'psychology' is the name of any unitary inquiry or tree of inquiries". "The facts which the psychologist claims to study are not different from those of biographers, historians, teachers, magistrates, politicians, or men in the street"; and psychology at best is just a "fortuitous federation of inquiries and techniques" (Ryle [45], pp. 172 and 319 f.).

In a paper published while these pages are going through the press (*Brit. J. Psychol.*, XLIX, pp. 106 f.), J. G. Taylor criticizes "the chaos of current psychological theory on grounds very similar to mine". I agree with his strictures, but not with his solution, which seems to turn on Lewin's false antithesis between Aristotelian and Galilean methodologies.

[2] I know of no actual analysis that answers the questions raised. Professor Braithwaite [4] has illustrated his own views of scientific method by inventing and analysing a miniature model, but he has undertaken no first hand analysis for any concrete science. Dr. Woodger [53] has endeavoured to construct a logical calculus for biology, and to do so has ingeniously adapted and developed the notation introduced in *Principia Mathematica*; but the result can hardly be regarded as describing the actual methods of the biologist, and the pages of symbolic proofs must be almost unintelligible except to those who have been trained in the symbolism employed. The nearest

In this paper I shall be concerned solely with logical structure in its most general form; and we may, I think, safely assume that the logical structure of every empirical science must in all essential features be the same. Which particular science we select, therefore, will not greatly matter. I shall choose dynamics, partly because it is at once the simplest and the most complete example of a science and consequently the one that comes nearest to a finished model or ideal, and partly because what may be called the dynamical aspects of psychology are the most neglected and the least developed. Whether psychology, more than any other science, contains something that cannot be fitted into the scheme that emerges I shall briefly discuss towards the end. In any case, a study of the thought-processes of the scientist in their most typical form should itself be a subject of interest to the student of psychology[1].

II. The Construction of a Scientific System

Choice of Model

Let us imagine that we wish to construct a scientific account of the behaviour, not of conscious or living organisms, but of unconscious material objects, and that we have resolved to begin with the simplest mode of behaviour—visible motion. Our first problem is to decide what kind of logical model we should adopt. The modern physicist commonly asks (i) what types of model are 'allowable' and (ii) which of the 'allowable' types is 'preferable'. There are two main alternatives, as old as the early Greek philosophers: the one treats 'the All' as a continuous but continually varying *plenum*, the other as a *vacuum* containing discrete but similar entities; the one is a 'field' theory, the other a 'particle' theory[2]. For the modern scientist both models are 'allowable';

approach to the analysis I have in mind is Pearson's outline of mechanics (with a couple of chapters on biology and statistics) set out in his *Grammar of Science* [34]. There, however, his main purpose is to advocate a new epistemology (or at least one that was new in his day, namely, that of Kirchhoff and Mach) in place of the materialistic assumptions of the nineteenth century physicist. As to method, the aim of science in his view is not to prove facts or theories, much less to offer explanations, but "to devise the most economical description of the phenomena encountered in sensory experience".

[1] I have learnt most from German and Austrian writers, particularly Hilbert [23], Cassirer [16], Weyl [51], Kraft [27], Planck [35], and Carnap [15]. For instructive discussions by English and American authors, partly based on the earlier German work, see [25] and [31]. These latter writers, however, are concerned chiefly with the metaphysical problems raised by the recent revolutions in modern physics (relativity and quantum mechanics); except Russell, they show little interest in the implications of their analysis for psychology. As will be seen from [43], pp. 141f., as compared with [44] Russell has greatly changed his views on this point; and I hope that we may say of him what the Prophet said of the Koran—that, where two texts seem inconsistent, the later should be taken as authoritative.

[2] The best account of the ancient controversy is Aristotle's (*De Gen. Corr.*, A, 8, 324b, 35; cf. *Phys.*, IV, 6). The 'double doctrine' of an empty space and discrete particles ('et corpora sunt et inane' says Lucretius) was developed by the early atomists (Leucippus and Democritus); Aristotelians, down to the time of Galileo and Gassendi, insisted on continuity and the *plenum*, and, like Nature as they understood it, 'abhorred' the notion of a *vacuum* as a mere container for particles.

and each has its own special brand of mathematical logic. The definitions of the leading concepts will inevitably differ in the two cases. Since the second model is the simpler and the more familiar, we shall select it here as 'preferable'.

Primary Concepts and Propositions

To make a start we must assume first of all a set of undefined concepts and secondly a set of unproved propositions: unless we do so, both our definitions and our proofs will be either circular or subject to an infinite regress[1]. We begin therefore rather like one of Lewis Carroll's 'Exercises in Versified Logic':

Let there be n unknown x's
Each with its own special a;
If a_i a b_i annexes,
What is x in terms of b_k?

But how are these primitive concepts and propositions to be chosen? Contrary to the assumptions made by most psychological writers, there is no decisive criterion for making a unique or optimum choice. 'Intuition' and 'self-evidence' prove highly untrustworthy guides. In geometry, for instance, several sets of postulates can be selected, besides those of Euclid, all of which yield much the same body of theorems: and the same holds true in empirical sciences. The only methodological requirements are that the sets themselves should be as elementary as possible, i.e. should readily lend themselves to further development[2]. The psychological requirements vary according to individual predilection and (so the historians assure us) different national traditions[3].

We start then with an assemblage of entities which, to begin with, are wholly undefined—M_1, M_2, etc., let us say. The symbols are just arbitrary monomarks, which distinguish the subjects of discussion, without telling us

[1] It may perhaps be suggested that those concepts which are not formally defined should be interpreted at the very outset by means of an 'ostensive' definition. And certainly I do not intend that the theoretical order adopted in my exposition should prescribe an order to be followed invariably in actual practice. But there are two potent reasons for postponing the interpretations. (i) As Hilbert [23] and the other 'axiomatizers' have argued, the only safe way to banish fallacies from our deductive arguments is to argue blindly, i.e. to "eliminate intuition wholly from the formal proofs" (a principle which many factorists might well take to heart). (ii) The undefined concepts and the propositions that are deduced will be found applicable, not only to quite diverse branches of empirical science, but also to different sets of objects within the same science, though only within a limited range. Thus in dynamics the undefined M's may in one problem be particles, in another cannon balls or railway engines, in a third planets or stars. (They cannot be protons, electrons, or neutrons, and it is doubtful whether they can be masses of gases or galaxies.)

[2] The more complex the subject, the larger will be the number of primitive concepts and propositions: in biology Woodger requires 10 undefined concepts and 35 unproved 'axioms' ([53], pp. 53 f.). The list of postulates given here is not of course intended to be complete.

[3] Cf. Duhem ([18], pp. 55f.). English and American physicists (and the ordinary man in the street) show a strong preference for what can be readily conceived in terms of a concrete visual and tactile image (e.g. quasi-mechanical models). Like Dr. Johnson, when they can kick a thing, they feel convinced that it is real; and when they can picture a process, they feel they understand it. German writers seem to work most easily with abstract algebraic symbols; the French are said to prefer a geometrical approach and lay special stress on aesthetic elegance.

anything about them, except that they are all to be treated as belonging to the same family or 'set'. To each M we shall assign four variable attributes, t, x, y, z, also for the present undefined—the barest minimum we can usefully take. Whereas the 'entities' are assumed to form a discrete set, each 'attribute' will be regarded as capable of continuous quantitative variation. And our first postulate (I) will be that the attributes conform to the axioms of a four-dimensional manifold.

In order to compare the entities with each other, we shall suppose that for one of the variable attributes, say t, the same value may be assigned simultaneously to all the M's. In order to distinguish them we shall assume that, when t has the same specific value for all the entities (t_0 say), only one of these 'entities' (M_1 say) can possess a particular group of values for the other three attributes, viz. x_1, y_1, and z_1. This restriction will constitute our second postulate (II).

We can now identify M_1 by saying that it is the M which, when $t = t_0$, has the attributes x_{10}, y_{10}, z_{10}; and similarly for each of the other M's. When t acquires a different value, say t_1 or t_2, then one or more of the attributes of M_1 may, or may not, acquire a different value, e.g. x_{11}, etc., or x_{12}, etc. To simplify the discussion we shall confine ourselves to variations in one attribute only, namely, x.

Derivative Concepts

The next step will be to construct certain derivative concepts, which will be wanted for the proofs that follow. These we shall build up by taking two or more of the undefined attributes, and combining them to make a new concept, which will thus be a specific 'function' of the old. The form of combination we shall indicate by signs called 'operators'. And it will be convenient to introduce a single symbol for each of the complex expressions so constructed. The proposition stating that the new symbol is equivalent to the whole complex expression (i.e. that one may always be substituted for the other and vice versa) will be called the 'definition' of that symbol.

In studying the changes in the attributes specified, we shall begin with an assumption that may be termed the 'differential postulate' (III). We shall assume that the expression used to describe each change will state in the first instance the momentary tendency at each *point* of the t-dimension (not the total outcome of that tendency after some finite interval); this means that the hypotheses we put forward will take the form of differential equations, with all that this procedure entails[1]. Accordingly, to specify the rate of change in x,

[1] In virtue of postulate III, no postulate of causation will be necessary; what were formerly regarded as causal laws will become differential equations, usually of the second order. The need for making the postulate explicit will be seen if we contrast the present model with that adopted in quantum theory. There the change is *not* expressed by differential equations, but treated as *discrete*, i.e. as a succession of states described first by one integer (or set of integers), and then by another; consequently, instead of determining the ultimate outcome by integrating along the route followed, the quantum theorist is compelled to state it in a statistical form, i.e. in terms of the proportion of cases in which each possibility is realized. The use of differential equations to describe mental changes (e.g. in growth and learning) has but recently been explored; and, not unexpectedly, non-mathematical psychologists have promptly criticized it along much the same lines as Berkeley adopted in his celebrated attack on Newton's use of the same procedure.

we introduce the symbol v, where $v = dx/dt$; and to specify the rate of change in this rate of change, we shall use the symbol a, where $a = dv/dt = d^2x/dt^2$. We could, of course, continue to build up rates of change in this fashion. But once again the methodological principle of maximum simplicity warns us to construct our model with the fewest possible components. We shall express this requirement here by two further postulates which must be regarded as provisional only, namely, (IV), that, so long as there is only *one* M in the field (M_1 say), the value of dx_1/dt, whether positive, negative, or zero, remains unchanged: i.e. $d^2x_1/dt^2 = 0$; and (V) that, when there are two M's in the field, then, if dx_1/dt changes at all, its rate of change, viz. d^2x_1/dt^2, will also be constant: i.e. $d^3x_1/dt^3 = 0$. This excludes derivatives of the third and higher orders.

Hypotheses

When we turn to fields in which *two* entities, M_1 and M_2 say, are present simultaneously, an obvious question at once confronts us. Suppose that during the interval $t_1 - t_2$ the x-value for M_1 changes from x_{11} to x_{21}, where x_{21} happens to be the value possessed by M_2 when t had the value t_1. Then, in virtue of postulate (I), the x-value for M_2 must also have changed. The rates of change in the two cases may of course be different. Accordingly, let v_1 and v_2 represent the values of v for M_1 and M_2 before t_2, and u_1 and u_2 the values after t_2. The problem then arises: what particular values will our model assign to these four symbols?

The aim of our whole enquiry is, wherever possible, to arrive at general rules or 'laws' whereby, if one set of n quantities are given, we can infer another set of m quantities ($m \leqslant n$) with a reasonable degree of accuracy. In the present case therefore we shall seek some means of expressing the second set of quantities (the u's say) as functions of the first (the v's). At this stage any such rule can be no more than a tentative surmise. To distinguish such conjectures from the postulates, I shall call them 'hypotheses'[1]. Here, in order to keep our deductive schemes as simple as possible, we shall assume that the function is linear, i.e. that the quantities concerned are additive; and our first hypothesis will accordingly take the form

$$v_1 + v_2 = u_1 + u_2. \qquad \text{(i)}$$

With this simple rule, when any three of the quantities are given, we can always deduce the fourth, viz. by means of a derived equation (the 'corollary') such as $u_2 = v_1 + v_2 - u_1$.

[1] I use the two terms in the sense defined by Johnson (*Logic*, I, p. 6). "A *hypothesis* is a proposition which awaits scientific investigation before being finally accepted or rejected: a *postulate* is a proposition . . . the truth of which is demanded by the thinker". The postulates in fact are the basic assumptions which together describe the model selected. (I have omitted Johnson's statement that postulates "cannot be brought to the test of experience", because they are in a sense tested indirectly when the model as a whole is compared with the factual data.)

Interpretation

In this elementary scheme we already have a miniature blue-print for a scientific theory. But it is a theory about a subject-matter that is so far wholly undefined. We might say of it what Bertrand Russell says of mathematics generally: it is a theory "in which we have no idea what we are talking about or whether what we are saying is true". It could apply to psychology as well as to physics. Before we can make practical use of it, we shall have to specify how our various concepts, both defined and undefined, are to be interpreted in terms of the phenomena we wish to study.

If, then, we are to test the corollary to which our hypothesis has just led us, we must first find some actual specimen—some accessible sample or 'instance'—of our schematic array of entities and attributes. In other words, we must state with what particular observables M_1, M_2, x, y, z and t are to be identified, and so build a bridge between our abstract symbols and the concrete facts.

We could, for example, identify M_1 and M_2 with a representative pair of *material objects*, say two balls on a billiard table: x, y, and z will then be the rectangular coordinates (referred to some convenient Euclidian frame of reference with a specified unit of measurement), and their values will jointly define the 'position' of each M in the *space* on or above the table: while t may be taken to denote mean solar *time* measured in seconds. Change in x, etc., will then mean change of position with change in time, or, in one word, 'motion'—the phenomenon we have set out to study; v will measure the rate at which position is changed. i.e. 'velocity', and a the rate of change in velocity, i.e. 'acceleration'[1]. We could, if we liked, call these interpretative specifications 'ostensive definitions' (in the case of the undefined concepts) or 'operational definitions' (in the case of the derived concepts); however, the danger of such phrases is that writers

[1] With this interpretation postulate II (p. 39) becomes equivalent to Newton's first 'law' of motion, which states that, until some further condition intervenes, each body continues to move with uniform velocity: (if, with the frame of reference used, that velocity is zero, the body continues in its apparent state of rest). To the plain man, as to the Aristotelians down to the time of Newton and Galileo, this sounds a wild paradox. Surely, it was said, a continuous pushing or pulling is necessary to keep things moving. At Oxford, however, Ockham (*c.* A.D. 1320) had long ago maintained that, when (for instance) an arrow leaves the bowstring and its motion continues till the air arrests it, there is "no need to imagine any further force or impetus"; and hence, in accordance with his principle of economy, "such a postulate is needless" (*Quaest. Lib. Phys.* VIII, qn. 12). The common statement that Galileo introduced the concept of 'acceleration' is quite mistaken: it was discussed by Aristotle. Galileo's novel suggestion was, so he asserts, that acceleration should be treated as *uniform*: "why should I not suppose that such increases take place in the simplest manner possible?" (*loc. cit.* p. 197). And this is provided by our postulate III.

May I here remind the reader that in dynamics the concept of 'acceleration' has proved far more helpful than that of either speed or velocity? Nevertheless, involving as it does the idea of a rate of rate, it strikes the plain man (at least until he learns to talk of 'accelerating' his car) as an absurd sophistication. That no doubt is why it so seldom enters into psychological discussions. The popularity of the I.Q., for example, rests on the intelligibility of a simple rate of growth, which is a velocity; yet in research it has proved worse than useless. On the other hand, as I have shown elsewhere, if we express processes like growth, learning, and fatigue in terms of positive and negative acceleration, we obtain a description much closer to the facts and far more accurate for purposes of prediction (see [10], pp. 664f. and this *Journal*, pp. 49–62).

who use them tend to overlook the arbitrary nature of the first group and the need of formal or constitutive definitions for the second. In point of fact, they are not so much 'definitions' as *ad hoc* instructions for translating our artificial and symbolic language into another language that is directly related to observable phenomena—i.e. into what is sometimes called the 'object language', "a language of the lowest possible type" (as the logicians quaintly put it), where the words can be understood without further elucidation in terms of other words.

Confirmation

(*a*) *With the Initial Problem.* Our hypothesis (eqn. (i)) can now be re-stated in simple verbal form: we may say that our tentative assumption is that the total quantity of motion (defined by such sums as $v_1 + v_2$) remains invariant, i.e. is 'conserved'. The alternative hypothesis is that this quantity is not 'conserved', e.g. that a head-on collision of two billiard balls would reduce it. To decide between the two we must plan a series of experiments to check the more obvious corollaries, such as that already deduced. And we then find that, to a tolerably close approximation, the values so predicted are correct.

(*b*) *With New Problems.* Now we come to a vital feature of scientific methodology which the psychologist too often overlooks. A hypothesis is rarely accepted just because it yields a satisfactory answer to the question with which the investigator started: the scientist nearly always expects it to have fresh applications as well—to suggest new problems and to furnish new solutions. In the present example we are led almost inevitably to ask: will the equation we have just confirmed hold good for other material objects, e.g. if M_1 and M_2 are taken to denote, not two ivory balls of equal weight, but (say) a ball of glass and a ball of lead? Further experiments reveal that it does not. In such a case the prudent course is not to reject out of hand a hypothesis that has already proved partly true, and so start all over again (as the psychologist so often does), but to reframe the hypothesis in a more general shape.

The next most simple formula will be an equation that includes one or two supplementary quantities (undefined at first) to serve as adjustable coefficients. Accordingly, eqn. (i) is amplified as follows:

$$m_1 v_1 + m_2 v_2 = m_1 u_1 + m_2 u_2, \qquad \text{(ii)}$$

where we assume that, for a given M (M_1 say), the value of the corresponding coefficient (m_1) remains constant no matter with which of our other M's we allow M_1 to interact; and similarly for the other coefficients. Experiments now show that corollaries deduced from this modified hypothesis apply with reasonable accuracy to a far wider range of phenomena. We therefore discard the simpler version, and accept the new.

It will be convenient to have verbal labels for the new expressions thus introduced. Let us call m_i the 'mass' of the material body, M_i, to which it belongs; and let us call $p_i = m_i v_i$ (which is, as it were, a weighted velocity)

its ' momentum '. Since p may also change its value with a change in time, we shall want a symbol and a name to express its rate of change. This we may designate F or ' force '. We then have

$$F = \frac{dp}{dt} = \frac{d(vm)}{dt} = v\frac{dm}{dt} + m\frac{dv}{dt} = m\frac{dv}{dt} = ma, \qquad \text{(iii)}$$

since m is constant and dm/dt therefore zero[1]. As we shall presently discover, F is a concept which, in its formal properties, exhibits a close resemblance to certain recent and rather elusive concepts occurring in psychology[2].

[1] As it stands this equation is equivalent to Newton's second ' law of motion '. If, however, we had started (as some writers do) by taking ' force ' as a basic concept, we could treat eqn. (iii) as covering both his first and his second ' laws '; for, when $F=0$, dv/dt must be 0, and whether v is also 0 or has some other value, its value will remain constant. The difficulty with Newton's formulation is that it is never clear whether his laws are intended as postulates, implicit definitions, or hypotheses needing confirmation. Hence, as has frequently been pointed out, the logical framework of his argument is decidedly lacking in rigour as judged by the standards of modern mathematics. The student interested in the conditions of valid empirical proof will find it instructive to examine the criticisms of Newton's logic advanced by Berkeley.

[2] Since attempts have been frequently made by advocates of a ' dynamic psychology ' to introduce the concept of force into psychology, it may be useful to examine this concept a little more closely. The physicists of Newton's day regarded force as a concrete causal factor of which we have direct intuitive knowledge through our experiences of muscular strain and resistance; and Newton himself almost certainly intended his second law of motion to supply, not (as some physicists have held) a novel definition of force in terms of momentum, but a synthetic proposition stating the relation between the two. In much the same way McDougall maintained that we acquire a direct acquaintance with ' mental forces ' through our experiences of conation. Nor do I see any reason why such experiences should not be refined and generalized, and so used as a basis for scientific concepts, or at least for interpreting the concepts provided by our models. If, however, we decide (as seems wisest in pure science) to dispense with the postulate of cause, then in physics, as we have seen, force must be treated simply as a convenient generic term for an algebraic expression, which continually recurs in our equations ($F = ma$), and which provides a useful middle term linking the general laws of motion with the special laws of nature. American psychologists (e.g. Woodworth in his *Dynamic Psychology*) seek to escape the ' animistic implications ' supposed to inhere in the popular term ' force ' by substituting the word ' drive '. To English ears the word sounds a little strained, though Americanizers have now led us almost to naturalize the usage. ' Drive ' in this sense is a Germanism; and in German *Trieb* denotes both ' instinct ' in psychology and ' impulse ' in physics: (in physics ' impulse ' is merely force multiplied by the time-interval during which the force acts, Ft, or the corresponding integral when the force is not constant: see p. 46 f.).

If we keep merely to formal characteristics, ' force ' in psychology would cover any kind of ' causal factor ' or ' potentiality '—' abilities ' as well as ' conative impulses '. Spearman preferred to describe ' abilities ' as ' energies ', i.e. as Fx rather than Ft; but surely ' time ' (t) is a more intelligible multiplier than ' distance ' (x). The formal similarities are tacitly implied by the terms used by Newton and by Galileo for ' force ': *vis*, *virtù*, *facoltà*; these are identical with the terms used by the scholastic philosophers to designate mental ' powers ' or ' faculties ', which played the same role in their physical theories as ' abilities ' or ' factors ' in modern psychology. And it is instructive to note that the Cartesian writers, who formed the orthodox school on the continent, brought against Newton's concept of gravitational ' force ' much the same objections as those urged by contemporary psychologists against the concept of innate ' ability ' (cf. Maddox [30]). An ' occult power ', it was argued, which vaulted over empty space with no material vehicle to transmit it, could not possibly be said to have real existence, and was therefore a purely ' metaphysical ' anomaly; and even *per impossibile*, supposing it to exist, it could never be directly measured or verified. Similarly in Galileo's dialogues, Simplicio (his scholastic critic) repeatedly objects: " How can a concept defined in a purely abstract manner correspond with what we meet in nature ? " (e.g. *Due Nuove Scienze*, *Giorn.* III, p. 198).

Vector Quantities and the Parallelogram Law

The coefficients m_1, etc., are necessarily positive, and are plainly scalar quantities. On the other hand, v, a, and F, being differentials derived directly or indirectly from the x's, y's, and z's, must have direction as well as magnitude: in short, they are vector-quantities. Consequently, when entering into formal proofs, they will call for the notation and the rules of vector or matrix algebra.

Our next step will be, once again, to pass from the simpler to the more complex problems, i.e. from the consideration of a single force to that of two or more forces operating jointly. What happens if we impart two simultaneous velocities (or accelerations) to M_1—one in the x-direction (v_x) and another in the y-direction (v_y)? To deal with this question it is simplest to introduce yet another postulate (VI), namely, that such quantities as v_x and v_y (and therefore a_x and a_y, F_x and F_y) combine in accordance with the rules for the addition of vectors, i.e. in accordance with what is sometimes called the 'parallelogram law'[1]. And here again we have encountered a type of situation which frequently occurs in psychology and which the psychologist can fruitfully handle by means of the same methodological device.

An Alternative Scheme

In the foregoing account we have introduced and defined the concept of 'mass' before making any reference to the concept of 'force' or 'weight'. But in point of fact this does not represent the way in which those concepts first made their appearance in the history of dynamics, nor the manner in which they are now commonly defined. And in order to illustrate how, within any branch of natural science, the definitions depend on the conceptual system selected, while the selection of the system often depends largely on convention, it may be worth while pausing for a moment to glance at this alternative mode of approach.

Classical dynamics sprang from Newton's interest in the 'fabric of the heavens' and in the movements of the celestial bodies isolated in 'outer space'—the sun, the earth, the other planets, and the newly discovered satellites. According to the story that Newton's step-niece told Voltaire, her uncle while

[1] This adds a new 'law of motion' to those ordinarily recognized—a procedure deliberately adopted by several modern writers (e.g. [50], p. 32). Alternatively, by substituting other postulates, it is quite feasible to treat the parallelogram law as a derivative theorem or a hypothesis to be tested: for this approach and a suggestive method of proof, see W. R. Hamilton, *Mathematical Papers* (1940 edition, II, p. 284). As applied to velocities, the formulation of the parallelogram law is as old as the pseudo-Aristotelian treatise (dating from about 250 B.C.) known as *Mechanica Problemata* (i and xxiii). Whichever procedure is adopted it is important to recognize that, like many other assumptions, the 'law' may involve a simplification that is not strictly warranted in practice. As Pearson (who was primarily responsible for introducing the device into psychology) is frank enough to avow, it rests on 'the tacit principle of the independence of causes', i.e. the assumption that, when two given forces appear in the same field, they leave each other unmodified: even in physics, he believes such an assumption may not always be legitimate (cf. [34], pp. 317f., 'On the Hypothesis of Modified Action and the Synthesis of Motion'). I take it that it is this assumption that Mr. Maddox wishes to repudiate when he speaks of a possible 'interaction of causes' in psychology [30].

meditating in the orchard at Woolsthorpe, was led to compare the relations between the earth and the revolving moon with those of the earth and an apple that had just fallen. Let us then begin, as he did, with the broad assumption that, when two such material bodies, no matter what their size or substance, appear together in the same spatial region, no matter where that region is—'on the earth below or in the skies above', they will affect each other's motion in a definite manner, whether or not there is any direct link or contact between them. And let us use a_{ij} to denote the change in M_i's motion produced by M_j, and a_{ji} to denote the change in M_j's motion produced by M_i. The simplest possible theory would be obtained if we could assume that the numerical values of a_{ij} and a_{ji} each remained the same, even when the bodies differed.

To put this to the test, let us imagine that M_1 and M_2 are two small spheres of the same size but of different material (e.g. wood and lead), and that the relation between the two is such as would arise when they were actually attached by an elastic band or a spiral spring. The experimenter pulls the two apart to a pre-assigned distance, releases them, and measures the resulting accelerations. They move towards each other. Hence their accelerations have opposite signs. But the rate at which the movement changes is different for each of the two bodies and with different pairs. Let us therefore introduce the symbol $m_{ij} = a_{ji}/a_{ij}$ (a positive scalar) to express the ratio of the two accelerations, and compare ratios instead of accelerations. The next simplest theory would be to suppose that, when the bodies M_1, M_2, M_3, etc., are allowed to interact in pairs, then $m_{ik} = m_{ij}/m_{jk}$, no matter to which body the suffix i refers. So long as we confine ourselves to macrophysical phenomena, there is nothing in the relevant observations or experiments that is grossly inconsistent with this assumption. Let us therefore adopt it as our next postulate (VII). It then follows that the ratios m_{ij}/m_{kj}, m_{il}/m_{kl}, etc., are all equal; and under these conditions we may validly select one of the bodies, M_k say, as a permanent standard. We can consequently drop the second suffix, and assign a single coefficient m_i to each body M_i. From the definitions of the m's it also follows that $m_i a_{ij} = -m_j a_{ji}$; and, if we put

$$m_i a_{ij} = F_{ij}, \tag{iv}$$

we then obtain $F_{ij} = -F_{ji}$, a theorem which is evidently equivalent to Newton's third law.

Equation (iv) may be regarded either as defining mass in terms of force or as defining force in terms of mass. We could therefore choose force as the more fundamental, as the engineer commonly does; and in that case we might use standard springs (placed horizontally) to provide units of force. However, in practice it is easier to preserve standard bodies of given mass. Hence the physicist generally prefers to treat mass as the more fundamental[1].

[1] With the change of conceptual system, the units change: thus, when force is the fundamental concept, the unit of force is the gram and the unit of the mass the metric slug; when force is the derivative concept, the unit is the dyne and the unit of mass the gram. Historically mass has entered into traditional mechanics in two ways—as inertial mass and as gravitational mass. Eötvos's

The Law of Gravitation

Now let us return to the case where no material link or spring connects the two bodies. In order to describe the motions of such objects as the earth and the moon or the apple in free space, we can adopt Newton's hypothesis in its original form, namely, that, when two bodies, of masses m_1 and m_2 respectively, are separated by a distance r, then the mutual acceleration which results is that which would be attributable to a force having the magnitude

$$F = (Gm_1 m_2)/r^2. \qquad \text{(v)}$$

Here G is a universal constant whose value has to be determined empirically[1]. For bodies near the earth the ratio Gm_1/r^2 (where m_1 is the mass of the earth) can be treated as everywhere the same and we can write it g. Then (v) reduces to $F_E = gm$, where $m = m_2$ denotes the mass of the smaller body. In this form, it will be observed, the equation implies that in a perfect vacuum all bodies fall

experiments with the torsional pendulum showed that the two were equivalent with a discrepancy of less than 1 in 10^8. Their absolute identity is, however, regarded as axiomatic rather than experimental.

The modified version of Newton's procedure given above is due to Mach [29], and in this country was first advocated (in a slightly altered form) by Mach's loyal disciple, Karl Pearson ([34], p. 302). It now appears in most textbooks. There is, however, a yet more general approach which to my mind is the most typical and suggestive of all. Having introduced our basic concepts, we write down our initial postulate in the most general form possible, $m(d^2x/dt^2) = f(x, dx/dt, t)$. To simplify the mathematics we begin by studying cases where f is a function of one variable only. Taking t first of all, after integrating both sides, we obtain $m(dx/dt) = S_1(t)+C_1$; and, after a second integration, $mx = S_2(t)+C_1\,t+C_2$, where the terms in S_1 and S_2 denote the integrals $\int f(t)\,dt$ and $-S_1(t)\,dt$ respectively, and C_1 and C_2 are constants of integration. This admirably illustrates how mathematical deduction may give rise to arbitrary constants which can only be evaluated from the empirical ' boundary conditions ' of the concrete problem. Writing $t = 0$, we see that C_1 is the initial velocity, v_0 say, and that $C_2 = mx_0$, where x_0 is the initial displacement from the arbitrary origin. If we further assume $f(t)$ is constant $= ma$ say, the ' equations of motion ' thus derived reduce to the familiar formulae (i) $v = at+v_0$, and (ii) $x = \frac{1}{2}at^2+v_0\,t+x_0$. (In passing it may be remarked that this generalized mode of approach has proved successful in dealing with the very different problems of mental growth and learning.)

We can proceed in the same way with the other variables, and then with two or more jointly. Here we have an excellent instance of a methodological procedure that can be applied to a wide range of problems in psychology as well as other sciences. Briefly it consists in picking out first of all the crucial variables and expressing the most complex of them explicitly as a generalized *function* of the remainder; then, to simplify the mathematical treatment, all of these except one are assumed in turn to be constant. The results are checked by empirical data; and supplementary variables reintroduced where necessary, by a process of successive approximation.

[1] It is in fact (in c.g.s. units) approximately $6{\cdot}67\times10^{-8}$. The earth's mass is $m_1 = 6{\cdot}0\times10^{24}$ kilograms ($5{\cdot}9\times10^{21}$ tons) and its radius $r = 6371$ kilometres (3965 miles). Hence $g = Gm_1/r = 981$ (in c.g.s. units) or 32·2 (in ft.-lb.-sec. units). Newton's verification of his law for ' the orb of the moon ' proceeds as follows: " If we imagine the moon deprived of all motion and let go so as to fall towards the earth, it would in the space of one minute fall $15\frac{1}{12}$ Paris feet. At the surface of the earth the force will be 60×60" (i.e. r^2) " times greater than at the moon" (since the distance of the moon is 60 times the radius of the earth); hence there a heavy body should descend $15\frac{1}{12}$ Paris feet (or 16·1 English feet) *in one second*. And this it does; for (as we have just seen) $x = \frac{1}{2}at^2$, where $a = 32{\cdot}2$ and $t = 1$. " And therefore ", he concludes, " the force whereby the moon is retained in its orbit is the same as the force of gravity we observe at the surface of the earth " (*Principia*, Bk. III, abridged).

to the ground with equal acceleration regardless of their masses—a paradoxical result that the mediaeval disciples of Aristotle found so hard to accept[1].

The force F_E as thus defined is technically termed 'weight'. But, since at a standard locality (e.g. sea-level, latitude 45°) g will be constant for all material bodies, 'weight' as thus measured also provides us with a convenient 'operational definition' for measuring their 'mass'. But this 'operational' procedure is quite unsuited to serve as a 'formal definition' of 'mass', since it yields no indication of the part which this concept plays in our conceptual scheme.

III. Hypotheses of a Higher Level

I have now, I trust, carried my outline far enough to illustrate the initial stages in the construction of a scientific model or scheme. Let us skip the next few stages, and consider how its development will proceed at higher and more complex levels.

Energy as Capacity for Work

When we go on to ask how the foregoing concepts will help us in discussing the everyday transactions of terrestrial life, we meet with problems which, in their formal character, present a much closer analogy with those which the psychologist has to consider. Human beings are called on to perform a variety of tasks varying widely in difficulty: these may be mental tasks, like answering test-questions, or physical tasks, like lifting weights. And, no matter what the task, there are upper limits to the difficulty with which different individuals can successfully cope. The same of course holds true of the machines that we employ as substitutes for human labour. In dealing with the issues thus raised—e.g. in trying to make theoretical predictions about the performance of some physical job—the physicist has recourse to a regular methodological principle: he endeavours to establish a *general law*, expressed in terms of some quantity that will remain *invariant*. It is worth while following the steps in some detail, since those psychologists who have borrowed the device seem to have overlooked the logical preliminaries.

Let us begin with a physical task similar to, but rather simpler than, that discussed in Professor Smith's first experiment—pushing a well-oiled trolley of mass m over a pair of frictionless rails through a distance x_1-x_0 during a time t_1-t_0. Putting F for the force required, and ignoring resistance due to air,

[1] The experiment of dropping balls of oak and lead from the leaning tower of Pisa was performed, not by Galileo, but by G. Coresio in 1612 and repeated by Renieri in 1641; both found that the heavier weight reached the ground first. Galileo (*Due Nuove Scienze*, 1638, *Giorn.* IV, 276: cf. *De Motu*, *c.* 1590) also notes the difference, but argues that it is (i) 'very small' and (ii) attributable to a 'retardation by the air'. He introduces the observation primarily to illustrate a methodological principle which he repeatedly stresses—the principle of simplification (see below, p. 49), viz., that, if we are to apply mathematical reasoning to observable phenomena, we must (as, he says, Archimedes did) ignore minor *perturbazioni* due to *accidenti* (irrelevant features), and proceed by 'abstraction', frankly admitting that "conclusions proved in the abstract may not be strictly true when applied in the concrete" (*ibid.*, 274–6).

we might consider either of the products Ft or Fx. The product Ft (impulse or thrust) is of special importance in gunnery and rocketry. For more commonplace problems the product Fx is of greater interest. To allow for possible variations in the force from point to point, let us express it in a more general form, viz. as the product–sum

$$\int_{x_0}^{x_1} F\,dx = U \quad \text{(say)}.$$

Then, since $F = ma$ and $a = dv/dt$, we have

$$U = m\int_{v_0}^{v_1} v\,dt = \tfrac{1}{2}mv_1^2 - \tfrac{1}{2}mv_0^2 = T_1 - T_0 \quad \text{(say)}. \qquad \text{(vi)}$$

If we call the quantity U the total 'work' done during the displacement of the truck from x_0 to x_1, and T the 'kinetic energy', we here reach an important theorem stating that "the work done on a body by a force is equal to the change in the kinetic energy that occurs during the displacement of the body".

Turn now to a different type of task where an appreciable resistance has to be overcome, say lifting a heavy body of mass m (a hammer or a pile-driver) through a height of x measured upwards from the ground. The force needed will be $F = mg$; and, since this remains constant, the work done will be $U = Fx = mgx$. Now let the hammer or pile-driver fall. Then d^2x/dt^2 (the acceleration) $= g$, dx/dt (the velocity on reaching the ground) $= gt$, and x (the distance to the ground) $= \tfrac{1}{2}gt^2$. Hence the kinetic energy,

$$T = \tfrac{1}{2}mv^2 = \tfrac{1}{2}m.2gx = mgx = U.$$

We may thus speak of the amount of work done by raising the body as having been stored, so to speak, within the body in the form of 'potential energy' (i.e. as energy due, not to motion, but to position), until the body is allowed to fall. Since, according to the formula we have used, the potential energy (V say) varies with the distance above the ground, while the kinetic energy varies similarly with proximity to the ground, we have $T + V = \text{const}$. We may, if we like, express this by saying that the potential energy is progressively 'transformed' into kinetic energy, in virtue of which the body becomes able to do work. This suggests the well-known definition (I should prefer to say description) of 'energy' as 'the capacity to do work'.

In the expression for energy ($\tfrac{1}{2}mv^2$), m is scalar, and v^2 a square which is necessarily positive and devoid of direction. We have thus succeeded in deriving a scalar quantity which remains invariant throughout the processes of mechanical motion. And we have proved the 'law of conservation of energy' as a theorem deduced from our definitions and postulates. As such it is valid only under certain limited and hypothetical conditions (namely, that the field of force is conservative), though, so long as a reasonable approximation to these conditions obtains, the 'law' can readily be verified.

The physicist now seeks to generalize this theorem into a universal hypothesis: he assumes that, even under conditions in which the proof is invalid, the law still holds good. To save it, he introduces a succession of supplementary

concepts—heat energy, deformation energy, the energy of electrostatic and magnetic fields, and so on. As thus amplified, the concept of energy becomes a kind of common denominator, linking together practically all the processes with which the physical sciences deal into one comprehensive quantitative scheme. Under various disguises it even figures in the transactions of modern commercial life: when we pay for our coal, gas, petrol, and electricity, and even our meat and bread, what we want and what we purchase is really energy stored in conveniently accessible forms. Plainly this is just the type of concept which we sorely need in psychology. And it is this type of concept which the psychological factorist, by methods that in many ways are formally analogous to those of the physicist, has endeavoured to supply[1].

The Treatment of Complex Problems

To make my main points as clear as possible, the illustrations I have hitherto chosen have been far more elementary than the problems with which the psychologist has usually to deal. Let us therefore enquire how the physicist himself attacks questions in dynamics that are too complex to be solved by the methods so far described. Three different expedients appear to have been tried: (i) reducing the complex problem to a much simpler form; (ii) generalizing the concepts found useful for the simpler problems so that they will apply to the more complex; (iii) generalizing the logical or the algebraic procedures.

(i) *The Simplification of Problems.* The first—a course that seems to the plain man to be anything but scientific, and, when adopted by the statistical psychologist, nearly always provokes criticisms from non-mathematical colleagues —may be called the 'precept of simplification'. It can be briefly formulated as follows: whenever a detailed model of the actual situation appears far too complicated for the available logical or mathematical techniques, substitute a

[1] Of this close analogy, as regards both methods and results, critics like Mr. Richmond, Dr. Heim, and Mr. Miles seem quite oblivious. Mr. Richmond protests against the "mathematical abstruseness" of factorial methods; Dr. Heim against the substitution of narrowly defined quantitative concepts for the more familiar concepts of the layman. But it will now, I hope, be obvious that, no matter whether our problem is physical or psychological, unless we can use mathematical models and formal definitions, rigorous demonstration becomes almost impossible.

In psychology the best known discussion of the concept of energy is that contained in Spearman's chapter on 'the hypothesis of mental energy' ([48], ch. ix). But his arguments confuse the formal character of 'energy', as defined in the mathematical model, with the concrete nature of 'energy', interpreted as the chemical energy supplied to the body in the shape of food, and converted by the body into mechanical energy, electrical energy, 'neural energy', and heat. When the factorist speaks of cognitive factors 'doing work', like 'forces', in virtue of the 'energy' they embody, he is merely referring to the formal similarity of the mathematical models used to describe mental 'work' and physical 'work'. When Spearman identifies the factor *g* with the "energy liberated in the nervous system by chemical change", he is employing an entirely different argument; and to suppose that this identification is confirmed by demonstrating the 'constancy of mental output in varying directions' is to mistake the whole character of the neural processes involved. The nervous system is in no sense 'a conservative field'; nor can its 'store of cerebral energy' or its 'output' be genuinely constant. Finally, when Spearman talks of *g* as a measure of mental 'energy', he really means a measure of 'power' (a criticism that he subsequently accepted).

simplified model which will at least yield an approximate solution. For instance, almost any function can be expanded in a power series; and under certain circumstances powers of a higher order can be neglected and the function treated as linear. Or again, almost every form of complex variation can be analysed into a sum of orthogonal components, diminishing progressively in magnitude; and the components responsible for the smaller contributions to the variance can usually be ignored. Aided by these and similar devices, as has very truly been said, "one of the greatest achievements of the physicist is his ability to make general statements about complicated systems without paying much attention to the details" ([50], p. 131)[1].

(ii) *Generalized Coordinates.* When dealing with a problem involving one or two particles only, we were concerned simply with movements of translation; and the positions could be specified by three rectangular coordinates in a three-dimensional space. When dealing with larger bodies formed by aggregates of particles, we must envisage their possible rotations as well. To cope with these more complicated problems Lagrange (1736–1813) introduced the notion of generalized coordinates, whose number would be equal to the number of degrees of freedom possessed by the whole system; this means representing the system by a model in multi-dimensional space. The method has formed the basis for nearly all subsequent work in theoretical dynamics. Expressed in matrix notation, for instance, Lagrange's equations have furnished the basis for the whole theory of aerodynamics, and formed the foundation on which modern quantum mechanics has been erected.

(iii) *Statistical Techniques.* However, when we come to the still more complex problems of a technology which seeks to substitute the heat-engine or the dynamo for human and animal labour, the logical models commonly used cease to be mathematically demonstrative, and become statistical. As a result,

[1] We have in fact already touched upon some of the oldest examples, namely, problems in which large, irregular, and heterogeneous bodies, like planets or projectiles, are treated as ' particles ' or ' mass-points '. But the most important cases arise in dealing with machines, whether artificial or natural (e.g. the human body). Oddly enough, it is the advocates of a mechanistic interpretation of the psychologist's problems—particularly physiologists and behaviourists—who most frequently warn him that, " until the geneticist has determined the detailed mechanisms by which cerebral characteristics are transmitted, and until the neurologist has discovered the neural structures which underlie conscious activities, it is useless, and indeed fallacious, to hope to reach trustworthy conclusions about mental capacities or mental processes ". Let us assume then, as I think we may, that for the practical psychologist every human being may be regarded in the first instance as essentially a machine. Now even the simplest machine, e.g. Archimedes' lever and fulcrum or Hero's pulley and cord, is always a *combination* of material parts; and, as those early Greek physicists and almost every engineer since their time have recognized, in order to describe the working of such a machine it would be as superflous as it is impracticable to trace the detailed working of each separate part: it suffices to determine the *resultants* of the forces applied or released, and of the various constraints imposed; and these, as a rule, may be readily inferred by comparing the work done *on* the machine with that done *by* the machine, in short, by comparing input with output and ignoring the details of the through-put. This is the simplification regularly adopted by the modern engineer when faced with what he likes to call a ' black box ' (cf. *Educ. Rev.* X, pp. 121f.). Thus, for the behaviourist to demand that all psychological problems must first be translated into terms of the underlying neurology is as though the physicist should insist that all problems of macrophysics must first be expressed in the equations of quantum mechanics.

the deductions no longer yield certain or necessary inferences, but only probable inferences, subject to appreciable error. In discussing the phenomena presented by gases, for instance, the attempt to cope with the behaviour of the various particles individually would lead to quite insoluble problems. The two most obvious alternatives are (*a*) to choose a 'phenomenological' model, describing the system in terms of its macrophysical characteristics and treating these as what they seem to be, namely, continuous variables (e.g. volume, temperature, pressure, and the like); this was the earlier mode of approach, and led to Boyle's law and ultimately to what is now known as 'thermodynamics'; (*b*) to treat the system as an aggregate of microphysical entities, viz. the constituent molecules, and merely determine the average of the various quantities that characterize them (e.g. their average velocity and their average energy): this has led to Maxwell's well-known law for the distribution of velocities (virtually the Gaussian), and to the rival versions of 'statistical mechanics'. Both methods offer instructive models for the analysis of the behaviour of human aggregates[1].

Coordinating Hypotheses

As the various realms of physical phenomena have been successively invaded and reduced to something like coherent order, still more ambitious efforts have been made to integrate the various hypotheses or 'laws', distinctive of each domain, into an all-embracing formula. A remarkable step was taken in the 18th century by Moreau de Maupertuis (1744) in a paper with the suggestive title *The Harmony of Different Laws of Nature which have Hitherto appeared Incompatible.* The 'Different Laws' were those governing the motion of particles and the propagation of light. De Maupertuis introduced a novel concept called 'quantity of action' (energy × time or momentum × distance); and his 'harmonization' consisted in showing how the different equations or 'laws' could all be compressed into the simple statement that the entities behaved in such a fashion as to "minimize the total amount of action"[2].

[1] Once again we encounter a contrast in attitudes analogous to the contrast between the preferences for a psychological and a neurological approach in discussing human behaviour. In the equations reached by method (*a*) time no longer appears as the independent variable, nor are the equations expressed in differential form: this limits the possibilities of the conceptual schemes, especially for temporal prediction. As a result, thermodynamics is commonly held to be the least satisfactory branch of physical science. Method (*b*) seeks to rectify these defects by starting with entities assumed to be closer to reality. This means new definitions for old terms. Pressure becomes mean force of molecular impacts; temperature becomes mean kinetic energy of the molecules; while increasing entropy is identified with increasing probability in the arrangement of the molecules —the most probable being the most disorderly or 'random'; and this in turn yields a practicable method for describing the temporal irreversibility of material processes.

[2] The writers of the Oxford school—Grosseteste (*De Lin is*, 75) and Roger Bacon (*Op. Maius*, IV, ii, 2)—held that "Nature operates in the shortest way possible"; and indeed the doctrine goes back to Hero and Aristotle (*Metaph.*, V, 6). In modern science theoretical models based on the principle of minimizing certain values (or maximizing their opposites) have played an increasing part. A familiar example is the way so many things—some would say all things—seem naturally to tend towards a state of equilibrium or stability. Under such names as 'homeostasis' or the like this is a

The equations of Lagrange were subsequently developed along these lines by Hamilton and Jacobi, treating 'trajectories' in dynamics on the analogy of the path followed by a ray of light. The outcome was a remarkably succinct formula expressing the total energy of the system as a 'Hamiltonian function' of the generalized coordinates and the generalized momenta[1]. This has paved the way for a 'harmonization' of classical mechanics and quantum mechanics with older equations of macrophysics. And so we can trace a succession of bold hypotheses, spreading in ever-widening circles, and aspiring eventually to encompass the whole expanding universe—an electromagnetic theory of light, a kinetic theory of matter, an electrical theory of matter, and finally an ingenious assortment of field theories intent on unifying the recalcitrant phenomena of gravitation and electricity.

Cybernetics

Finally, let us notice, this interchange of concepts and methods between physics and psychology is by no means a one-way traffic. The original implications of terms like 'force' and 'work' were manifestly anthropomorphic. De Maupertuis's own interpretation of his principle was frankly teleological. When we come to the theory of modern man-made machines, the language of the physicist and the engineer, paradoxically enough, becomes less and less mechanical, and acquires a curiously psychological ring. We hear of 'goals' and 'goal-seeking tendencies', of 'behaviour' based on 'information', of 'input', 'output', and 'feed-back', of 'communication' more or less impeded by 'noise'. The notion of 'rational order', contrasted with 'randomness', 'disorder', and 'chaos', comes more and more to the front. Cybernetics itself is defined as "the science of communication and control in the animal and the machine" [52]; and the methods of matrix algebra that have been elaborated to deal with its problems betray a striking likeness to those of the factorist

concept exploited by physiologists like Claude Bernard and neurologists like Cannon, by philosophers like Avenarius and psychologists like Stout [49]. In a highly original treatise entitled *Mathematical Physics* (too little known to students of psychology) Edgeworth once proposed to construct a comprehensive calculus of human behaviour, based on the assumption that the behaviour followed in any situation would approximate to that which would maximize 'utility' (which for a 'utilitarian' meant maximizing 'happiness'). As a methodological device the principle of maximization was deliberately adopted in the earliest forms of factor analysis, where a canonical series of 'factors' are extracted by maximizing in turn the successive factor variances [11].

[1] These newer developments have proved capable of application, not only to the problems of classical dynamics, but also to those of relativity and quantum mechanics. As De Broglie first pointed out, the equations of Hamilton and Jacobi "retain their significance in quantum theory, whereas those of Lagrange do not"; and in De Broglie's hands they were made to prepare the transition from the older mechanics to the newer so called 'wave mechanics' ([9], pp. 34f.). Eddington, in his entertaining chapter describing how the scientist sets about 'World Building', suggests an explanation: "Hamiltonian differentiation is the symbol for creating an active world out of a formless background: it covers most of the operations needed in constructing the physical world" ([19], p. 240–1). In cybernetics 'information' takes the place of energy; but the model describing its transmission, direction, and conservation is much the same. On the applicability of the methodology of quantum mechanics to psychological problems, see below pp. 77–93.

(cf. [1]). However, with these newer developments in methodology I have dealt more fully elsewhere, and here need say no more[1].

IV. The General Structure of a Scientific Conceptual Scheme

The Completed Scheme

From this brief survey let us see how far we can piece together the general nature of the methodological scheme which forms the structural framework of any successful science. If we may take dynamics as the nearest approximation to the scientist's ideal, it would appear that the finished system should exhibit a hierarchical structure, in which levels of increasing generality lie one on top of the other, like courses of overlapping bricks. It would consist essentially of (*a*) a hierarchical classification of increasingly complex concepts and (*b*) a hierarchical arrangement of increasingly general propositions, linked into a consistent logical whole by what may be called 'implicative' relations[2], thus forming (in Bacon's picturesque phrase) 'a soaring pyramid of knowledge'. But it is a pyramid of knowledge resting on a platform of abstractions.

Working backwards, the whole procedure is rather like that immortalized in the rhyme of the House that Jack Built: "This is the generalization, derived from the laws, deduced from the equations, relating the *f*'s with the *m*'s and the *a*'s, ..." and so on, until we get down to the highly abstract symbols, *x*, *y*, *z*, and *t*, specifying the dimensions of the 'manifold', and lying at the very base. Working upwards, we start with an artificial groundwork of undefined concepts and unproved propositions or postulates. On this foundation are erected, first, a number of derivative concepts of increasing complexity, formally defined in terms of the primitive concepts, and secondly a succession of derivative theorems, formally inferred from the definitions and postulates, and culminating in the final formula. The theorems, however, must first be tested by planned observations or experiments, before they can be fitted firmly into place. Gaps are subsequently filled in, wherever possible, by the same critical process—trial, elimination, and survival of the fittest[3].

[1] *Educ. Rev.* X, pp. 117f., and this *Journal*, IV, pp. 195–200: cf. also T. F. Ramsay, *ibid.*, X, pp. 126–7.

[2] I use this term to designate what older logicians called the relation of 'ground and consequence' (commonly expressed by such conjunctions as 'if ... then', 'because ... therefore'), and the plural to cover such relations as 'the proposition *p entails* the proposition *q*', 'the set (or class) A *is included in* the set (or class) B', 'the property (or relation) K *involves* the property (or relation) *Y*'. The ensuing scheme might be succinctly described as a 'deductive' system; but it seems better to avoid this epithet since it is apt to be contrasted with, and so taken to exclude, induction.

[3] The common statement that the first attempts at a genuine scientific method along these lines were due to the genius of Galileo is far from correct. Euclid's remarkable achievement in arranging geometrical knowledge in a *sequence* (the title στοιχεῖα, mistranslated 'Elements', means 'the components of a series') supplied the original model for purely deductive systems. Archimedes and the Alexandrian scientists showed that the same postulational procedure greatly facilitated the discussion of problems arising in the field of natural science. And the combination of Aristotelian logic with geometrical formulations and experimental tests—so characteristic of Galileo's dialogues—had already been anticipated in the writings of Grosseteste, Roger Bacon, Ockham, and other members of the Oxford school. Roger Bacon, for example, in pt. vi of his *Opus Maius*—entitled 'De Scientia

In any particular section it is the narrower or more specialized hypotheses that are commonly laid down first (e.g. Kepler's final verification of Arzachel's conjecture—that the planet Mars might move in an elliptical and not in a circular orbit). These correspond with what Mill calls '*empirical laws*, based on direct observations', and Whewell 'laws of phenomena'[1]. Broader and more comprehensive hypotheses are then constructed, which would logically entail, and so coordinate, the more limited hypotheses (e.g. Newton's demonstration that a central force, inversely proportional to the square of the distance, will furnish the equations describing not only the elliptical orbits of Mars and the other planets, but also the orbits of the satellites round Jupiter and the moon round the earth, the parabolic motion of comets, and so on). These broader hypotheses correspond to what Bacon called *axiomata media*, Mill '*intermediate principles*', and Whewell 'causal theories' (on the ground that they 'explained descriptive hypotheses by causal concepts'). Finally, an endeavour is made to integrate these mid-level hypotheses into a single all-embracing 'theory'—what Mill calls the '*highest generalization*' (e.g. Newton's universal law of gravitation, eqn. (v) above, from which we can deduce both the motions of heavenly bodies in the sky and Galileo's laws for falling bodies and 'military projectiles' near the surface of the earth). The fruitfulness of this unifying theory is then exhibited by demonstrating its applicability to entirely different problems—the motion of pendulums, the flux and reflux of the tides, the precession of the equinoxes, and the spheroidal shape of the earth. And the whole interlocking system thus constitutes what Jevons described as a 'hierarchy of natural laws' ([26], pp. 742 f).

System and Lack of System in Psychology

How far does modern psychology conform to this conception of a science? It might, I think, fairly be said that the leading psychologists of the later nineteenth and earlier twentieth century—say roughly from the first edition of Wundt's *Grundzüge* (1874) to the last edition of Stout's *Manual* (1938)—seem to have had in mind an ideal or a model broadly resembling what I have described. They attempted to draw up a systematic hierarchy of concepts: the isolated

Experimentali'—lays down a theory of natural science based on this combination and then (in that work and elsewhere) applies his methodological theory in several scientific researches, undertaken as much to illustrate the method as for their own interest. Galileo's logical terminology is derived partly from Aristotle but mainly (together with the orderly method of proof) from Euclid: 'definitions', 'axioms', 'theorems', 'hypotheses' (*argomenti ex suppositione*), and 'corollaries' are usually labelled as such, much as they are in scientific writings of the Alexandrian school. Most of his experiments are *Gedankenexperimente*; and he still holds that "without depending on experiment, which doubtless is decisive" (in refuting critics), "one ought to be able to establish the fact by reasoning alone". The idea that in empirical science *all* hypotheses should be checked by actual experiment is not found until we come to the works of Robert Boyle, who in this respect seems to have greatly influenced Newton and other leading members of the Royal Society.

[1] Mill, J. S., *A System of Logic*, 1856, II, pp. 36f., 447f.; Whewell, W., *Novum Organon Renovatum*, 1858, pp. 118f. Eddington and others distinguish between 'primary' and 'secondary laws', though the mode of distinction varies considerably ([19], pp. 75f.).

'sensation' plays the role of a single particle, and is treated as possessing four main aspects or 'dimensions'; the fusion of two sensations (one 'peripherally excited' and the other 'centrally excited') yields the simplest type of 'percept'; and the more complex aggregates—the 'concepts'—are classified under somewhat conventional headings, each methodically defined. There is also an attempt, far less successful, to construct a hierarchy of 'laws', and to deduce them from some cardinal and comprehensive principle—the 'law of association' with the earlier authors, a law of organization or 'noetic synthesis' with the critics of associationism[1].

Later writers, particularly in America, evidently felt that a procedure of this kind would be exposed to the familiar accusations of 'psychological atomism', 'premature systematization', and 'loss of touch with real life'. Hence more recent textbooks commonly abandon the attempt to integrate the facts and theories they describe into anything like a logically coherent scheme[2].

V. Recent Methodological Proposals

As we saw at the outset, several attempts have recently been made to "improve the scientific status of psychology by paying greater attention to scientific method"; and the foregoing outline of the methodology of a typical science should now enable us to assess the merits of the proposals thus advanced.

[1] See Murphy's account of 'how psychologists conceived the fundamental task of psychology as a science' ([33], pp. 213f.). The earlier associationists and their critics, while accepting physical science (including chemistry) as their model, adopted an excessively *statical* standpoint; McDougall's emphasis on conation, and his theory of 'mental forces' and a central conative energy, together with the investigations of educationists and industrialists on 'mental work', introduced a definitely *dynamic* standpoint: the psychiatrists frequently borrow terms from dynamics—'strain', 'stress', 'tension', 'reaction', 'resistance'—but use them in a loose and ill-defined sense. Spearman's theory of noegenesis, with a central cognitive 'energy', three 'primary laws' (framed on the analogy of Newton's three laws), and five 'secondary laws', yields hypotheses of the intermediate and highest level; but his deduction of hypotheses at the lowest level, for the purpose of experimental corroboration, is decidedly inadequate.

[2] Thus Woodworth's textbook (one of the first to set the new trend) deals almost from the start with complex and 'realistic' problems ('intelligence', 'personality', and the like), and postpones more technical matters such as the structure of the nervous system and the elementary processes of sensations to later chapters. Boring, Langfeld, and Weld explain in their preface that they have largely reversed the 'conventional arrangement' ('sensation at the beginning and thought and personality at the close'). Their original intention (they say) was to substitute 'analysis' for 'synthesis'; but in the end they "confess to a certain dependence on trial and error"—not, as might have been expected, on an order dictated by purely logical considerations. Moreover, whereas the older writers gave formal definitions for all their key concepts, the newer generally aim at popular explanations or picturesque descriptions, which often leave the precise scientific meaning blurred and ambiguous. For purposes of pedagogy and elementary exposition this modern approach is doubtless more effective; but as a method of scientific demonstration and training it seems, not merely inadequate, but at times bewildering. McDougall, it may be observed, attributed the change to the fact that "in America teachers of psychology" (and he included himself among them) "find by experience that in lecturing to, and writing for, the rapidly growing number of students of psychology, who are for the most part interested more in practical work than in theory, the only feasible policy is to put the easiest, most interesting, and most popular topics first". And he believed that he detected "a similar bias over here".

1. To begin with, however, we must face the prior objection that the problems of the psychologist are from their very nature unsuited to investigation by the methods of science. There are two recurring arguments: they have, however, been discussed so frequently that I can deal with them quite summarily here[1]. (i) First, it is argued, the subject of psychology is not 'mind as such'; for there is no such thing. "Its only possible subjects are the 'individual personalities'—Tom, Dick, and Harry." Moreover, each personality constitutes a 'unique whole'. Hence the most the psychologist can do is to seek to 'understand' it in the light of direct experience and intuition; he cannot possibly 'explain' it by any kind of scientific analysis. This seems rather like arguing that a really scientific study of the solar system is out of the question, because all that the astronomer can observe are the individual planets—Mercury, Venus, Mars, and the rest, each constituting a 'unique whole'. But individuals may have general qualities; and wholes may be complex and therefore susceptible of analysis. Intuition itself is only a form of observation; and, as with all observations, the inferences based on it must be expressed in general terms and must be subject to verification. My answer therefore would be that in psychology both the intuitive and the analytic modes of approach are possible; both have their uses; and the one need not preclude the other. (ii) But, it is objected, "the methods of science are essentially quantitative, i.e. numerical. And psychological phenomena are not amenable to numerical treatment. Hence psychology cannot constitute a genuine science: rather it is an art like medicine or a narrative like history". Here, however, the major premiss is too sweeping. In many sciences (e.g. botany, physiology, and even physics) much progress has been achieved, particularly during the earlier stages, by methods that are purely observational and qualitative. Moreover, qualitative attributes, no less than quantitative measurements, may be handled effectively by mathematical devices—a point that the critic nearly always forgets. When we look back over the instances of scientific reasoning given in the foregoing survey, it will be seen that numerical quantities play a comparatively minor role. The main value of a mathematical treatment springs more from its use of symbols than from its use of numbers; by inventing its own notation it greatly enhances the exactitude and the certitude of the inferences drawn. To rely on verbal arguments instead of on mathematical, as the intuitionist would have us do, is to suppose that exact demonstrations can be achieved with inexact words. Still, be that as it may, whatever procedure we resolve to adopt, whether mathematical or non-mathematical, we must at all events observe the basic requirements of logical

[1] Those who desire a fuller statement of the *pros* and *cons* may refer to the 'Symposium on Personality' (*Brit. J. educ. Psychol.*. XVI, pp. 5f., XVII, pp. 1f.): Dr. Maberly advocates a purely intuitive 'understanding'. He is answered by Allport (*ibid.*, XVI, p. 57; cf. *Personality*, ch. xix). And my own views are stated in the final contribution. The contrast between a *verstehende* and an *erklärende Psychologie* is due chiefly to German writers—Dilthey, Spranger, Jaspers, and other pioneers of 'existentialism'. The argument that psychology is not amenable to quantitative treatment and therefore cannot constitute an empirical science originates with Kant.

method; otherwise our conclusions can carry no conviction. It is logic, not mathematics, that is essential to a science[1].

2. Of those who accept the problems of psychology as a legitimate subject for scientific study, the majority, as we have seen, make the principle of an 'operational test' the crucial feature of their methodology. And this, they hold, implies that "all concepts and propositions which cannot be operationally defined must be discarded"[2]. Our review of the procedures actually adopted in science will now have made it clear that this inverts the true procedure: the theory of the concept is prior to the operation for measuring the concept, not

[1] Possibly all that is really intended by those who stress the second of these contentions is that psychology is not a science in *their* sense of the word, i.e. what they would call an 'exact science'. "But even if", they may say, "you insist on using the word in a broader sense, that would not alter the essential disparity: the subject matter of psychology is so different from that of the older sciences, and its problems are so much newer and more complex, that no analysis of the methods adopted in dynamics can possibly serve as a guide. Why not choose something equally novel, like quantum physics, or nearly as complex, like biology?" To this my answer would be that these branches of science are still too incomplete and controversial to reveal the kind of logical structure at which the scientist should aim. Moreover, such a choice would require greater familiarity with technical details than most students of psychology possess. The critic, however, who prefers these alternative models will find the methodology of the biological sciences described in precise symbolic notation by Woodger [53] and a discussion of quantum-theory in the later pages of this issue (pp. 79–93).

But to my mind the common notion that the problems of physics are characterized by their extreme simplicity is profoundly mistaken: it is presumably due to the fact that the elementary textbook seldom goes further than the simplified experiments of the laboratory. To predict the effects of the physical disturbances we meet in nature—a thunderstorm or a conflagration—is as tricky as predicting the effects of a mental disturbance on this or that individual. In my view the main divergence between psychology and physics springs, not from differences in complexity, but rather from the peculiar fact that the psychologist has the privilege of inside information about one particular specimen of the entities he studies: he can observe his own mental processes from within. For a behaviourist such a difference of course does not count; and at most it can only mean that the *detailed* methods of psychology (as of every science) must in some measure be peculiar to itself—a truism that has no relevance here. The *general* logical structure of every science must, as I stated at the outset, remain the same, no matter what its subject may be.

As a postscript may I add that, in a volume published since the above was written, this view has been cogently argued by Professor Popper in his discussion of the 'unity of scientific method' ([39], sect. 29)? To drive home his point he shows how the account of the methodology appropriate to a social science, as described by such writers as von Hayek (*Scientism and Society*), would serve equally well for any of the physical sciences. "All sciences", he concludes, "make use of the same logical method, whether they are natural sciences or social sciences" (and among social sciences he expressly includes psychology). And he too registers a protest against "the widely held prejudice that social situations are more complex than physical" in terms far stronger than I have used.

[2] It is not obvious how an operational definition should itself be defined. The most satisfactory definition is that given by Carnap (*Phil. Sci.*, 1946, III, pp. 419f.), which he himself apparently has now abandoned. Most psychologists rely on Bridgman's earlier account [5]. This, however, is neither clear nor self-consistent. Bridgman at that time had chiefly in mind Einstein's doctrine of relativity; and the 'operations' envisaged were essentially instrumental or experimental. However, such operations are often unsuitable for the concepts needed by the more recent developments of the quantum theory; and Bridgman himself is now willing to include purely 'mental operations'. In that case the requirement seems robbed of its novelty, since every type of definition involves a 'mental operation'. In psychology current discussions of the problem have been greatly influenced by the view held by many logical positivists—that methodology is essentially a critique of the scientist's language. To the scientist, however, what is important is, not so much the language, as the concepts that his language seeks to express, and the functions those concepts are intended to fulfil in building up his theories.

vice versa. Moreover, the injunction could not be consistently obeyed even in those sciences for which it was first suggested. As we have seen, no 'operational definitions' of mass or of force are needed until we require to measure them; and even then the so-called 'operational definitions' are not true definitions at all. Examples could easily be multiplied. What operational definition, for example, could be given of a 'neutrino'[1], or of Schrödinger's 'wave function', or of 'energy' in general as distinct from specific kinds of energy?

But my main objection to the doctrine is the common assumption that this is the only kind of definition needed or permissible. The very phrase clearly betrays a confusion between two distinct methodological steps: (i) relating the *definiens* to the *definiendum*, and (ii) relating the *definiendum* to what may be called the *referendum*. As we have repeatedly seen, what are really essential for the logical proofs are what are sometimes called the 'formal' or 'constitutive' definitions: and these must depend on, and often vary with, the particular conceptual scheme of which the concept forms a part. Let the reader take the 'operational definitions' for a set of concepts, either in physics or psychology (together with the usual postulates), and see how little can be logically deduced for purposes of empirical verification, unless we covertly drag in a constitutive definition by the back door[2].

3. More recent critics have fixed their attention, less on the way the psychologist constructs or defines his concepts, and more on the manner in which he seeks to establish his final conclusions. As we saw at the outset, there are two opposing doctrines, based ultimately on the traditional contrast between 'inductive' procedures and 'deductive'.

One group of writers, reviving the precepts of Francis Bacon, insists that in an empirical science the correct procedure is to start by collecting, tabulating, and classifying facts, and "to allow the hypotheses to emerge directly out of the observations" (cf. p. 35 sup.). If this means that we should keep what is observed distinct from what is inferred, it is wise advice. If it means that, as

[1] A neutrino is an elementary particle having a spin of $\frac{1}{2}(h/2\pi)$, but no mass (or practically none), and no electric charge, and hence can interact with no known observable.

[2] Since this further contention is so generally overlooked, one or two illustrations may be helpful. Take one from a deductive science first of all: suppose Euclid had taken the 'operation' described in Bk. I, Prop. 1, as giving a 'definition' of an equilateral triangle; that would have played havoc with his systematic classification of triangles and rectangles, and with all his subsequent proofs. Or again suppose that Galileo had substituted his operational definition of a parabola (*op. cit.*, *Giorn.* II, p. 149) for the usual formal definition: it would have been almost impossible for him to show that the trajectory of a projectile was a parabola, or for Newton to show that such a curve could be deduced from the law of inverse squares, or for the modern physicist to generalize this to the more correct statement that the trajectory is an ellipse (a correction ignored in most textbooks)—particularly as Galileo's 'operational definition' described not a parabola but a catenary! In psychology the favourite case (cited by Mr. Miles) is the definition of intelligence as "what intelligence tests measure". Suppose then that speed were defined as "what speedometers measure", force as "what dynamometers (calibrated springs) measure", and mass as "number of grams recorded by a weighing machine" (all of course under standard conditions); could we logically derive the equations of dynamics from these definitions? And, if we found errors in the experimental results gathered to check those equations, how should we know whether the errors should be ascribed to the inaccuracies of the theory or to the inaccuracies of our instruments and operations?

a matter of personal history, hypotheses have frequently been suggested by observations, it is unquestionably true. But most of those who hold this view take it to imply that "the appropriate method for an inductive science is the inductive method"; and induction they consider to be "the opposite or inverse of deduction" ([26], p. 122). "In the deductive sciences we start with general postulates or hypotheses, and deduce particular propositions; in the inductive we start with particular propositions, based on observation, and from them infer general hypotheses or principles."

Now, to put this forward as a general methodological precept for those embarking on a scientific research is, to my mind, gravely misleading. It is making the cart draw the horse. When a new enquiry is being planned, the hypothesis to be tested should always be explicitly formulated before any systematic observations are undertaken. Only after the crucial issues have been defined and the possible alternatives reviewed—and, as a rule, only after the distinctive implications of each have been logically deduced—is the investigator in a position to say what observations will be relevant, or how a conclusive experiment is to be designed[1].

4. At the moment the most fashionable view maintains that, in the natural sciences and therefore in psychology, the only correct procedure is the 'hypothetico-deductive method'. It will now, however, be clear that, as ordinarily described, this view yields a very inaccurate picture of the procedure actually adopted by scientists in the fields we have examined. To begin with, it entirely ignores the fundamental distinction between postulates and hypotheses.

[1] In my view, Poincaré [36, 37] and Duhem [18] have clearly shown how impossible it is to regard the theories of the so-called inductive sciences as inductive generalizations; and their conclusions, I fancy, are amply borne out by the examples given in the preceding survey. For the English student perhaps the clearest statement of the real difference between the deductive and the inductive method is that given by Joseph (*Introduction to Logic*, pp. 396–7).

I admit that to argue about the priority of observations and hypotheses is rather like asking which comes first, the egg or the hen. With any individual problem, the points to be observed must develop out of the hypothesis, just as with any individual hen the feathered bird is hatched out of the shell: but every hypothesis, like every egg, has a long ancestral history behind it. However, Mr. Jackson, who has been good enough to read the above in proof, puts it to me that my "own account of the development of the 'new quantum theory' (pp. 77f.) seems to show that, for the construction of a successful hypothesis, the collection of a vast array of facts is needed first—as with Rutherford's discovery of alpha particles, Planck's discovery of the quantum, Einstein's discovery of the photon, or Bohr's discovery of energy levels". My answer is that Mr. Jackson, like many other supporters of this view of scientific method, is confusing low level hypotheses, confirmed by previous observation, with 'facts already observed'. Rutherford did not 'discover' alpha particles: he pointed out that two hypotheses were conceivable to explain the nature of the radiations emitted by uranium—the 'rays' might be either charged particles or electromagnetic waves; in the former case they would be deflected by an electromagnetic field (the deflection varying according to the nature of the charge and the mass); in the latter case they would be undeflected, like x-rays. His experiments showed that certain of the radiations *were* deflected, and in such a manner as to indicate that they had a positive charge and a comparatively large mass: these were the 'alpha' rays. Similarly with the other 'facts' Mr. Jackson mentions. I do not deny that an accidental discovery of a 'fact' may sometimes come first—as with Becquerel's observation that salts of uranium blacken a photographic plate. But what such an accidental discovery usually suggests is a problem rather than a hypothesis. And with rare exceptions most 'facts' are 'discovered' because the investigator was deliberately looking for something of the sort.

As a result, its advocates continually mistake what are merely methodological conditions for conjectures requiring empirical proof. There is all the difference in the world between thinking of intelligence *as* a general factor and thinking that intelligence *is* a general factor; and the same phenomena may be looked at or described from very different angles, each defined by its own set of postulates: seldom, if ever, do the facts allow of only one interpretation. Thus factor analysis is constantly criticized because it fails to provide a unique solution[1]. This is rather like criticizing map-making because it provides no unique procedure for representing the surface of a sphere on a two-dimensional plane: a glance at a schoolboy's atlas will show that the cartographer is able to make effective use of several different modes of projection—conical, globular, orthographic, or Mercator's, according to the aim he has in mind. In precisely the same way the psychologist selects different types of model, varying with his purpose at the moment; and his postulates merely describe the particular model he has selected. Hence the function of a postulate is quite different from that of a hypothesis.

But even if we waive this general criticism, it is clear that the hypothetico-deductive method as ordinarily stated is fallacious in itself and inconclusive in actual practice[2]. It implies that, in solving any particular problem, only one hypothesis calls for testing; whereas two at the very least are essential, namely, (i) the hypothesis to be established and (ii) either its opposite or the next most plausible alternative. The proof then consists, not merely in confirming (i), but in refuting (ii). And, if it is to carry any conviction, such a proof implies that, before the observations were actually made, it must have been quite conceivable that the results obtained would possibly refute (i) and confirm (ii).

[1] " Thurstone ", says Richardson, " analyses the correlations between a set of temperamental assessments into a series of positive oblique factors; Burt analyses the same table into a general factor and a set of bipolars. Both cannot be true. Hence the whole factorial method stands condemned; and the proper verdict is that psychologists have so far failed to discover the true facts ". This is like complaining that, when two astronomers analyse the same set of data in terms of rectangular and polar coordinates respectively, both results cannot be true, and hence the whole method of coordinates stands condemned. As our review has shown, science does not by any means consist in merely " discovering facts ". Some of its most fruitful results are discoveries not of new facts but of new ways of representing the facts, which in turn give rise to new techniques for analysing and arranging the facts and drawing valid inferences.

The peculiar status of a postulate is so frequently overlooked that perhaps one or two further instances are needed to make it clear. In physics, according to the problem to be solved, the propagation of light may be described either in terms of ' rays ' travelling in straight lines (no one ever ' discovered ' rays or asks whether a ' ray ' exists), or in terms of electromagnetic ' waves ', or in terms of particles or ' quasi-particles ': electricity may be treated as a fluid, as two antithetical substances, as a gas-like collection of corpuscles, or in terms of a field. It is, I think, true that at times what in one model appears as a postulate may in another be treated as a hypothesis: e.g. Bohr's model of the atom as a solar system or again the law of ' conservation of energy ', each of which was originally put forward as a postulate, and both of which, when broader postulational models were devised, became hypotheses, and were experimentally disproved, at least in their original form. But within one and the same postulational system the distinction between postulate and hypothesis seems clear-cut.

[2] The fallacy has frequently been exposed: cf. Burt ([12], 1940, pp. 33f.) and Reichenbach's criticisms of the hypothetico-deductive method, *Theory of Probability*, 1949, pp. 431f.

In either case it is the rebuttal of one or other alternative that is really vital, and with some problems may be all that is feasible.

"The oracles of Delphi", says Cicero, "were so deftly worded that in either event they proved true"; and the stories of Ahab, Croesus, Philip of Macedon, and Maxentius remind us how alive the ancients were to the fallacies of equivocal predictions. It is just here, I fancy, that we may discern the chief difference between a pseudo-science and a genuine science. Astrology, for example, presents an ideal hypothetico-deductive system. The vague corollaries deduced from its hypotheses can always be confirmed. But its votaries make no attempt to disprove alternative hypotheses; and its doctrines contain so many loopholes that its own disproof is almost impossible. The same is true of many psychological theories[1]. The psychologist is continually warned that, if his hypotheses are to be meaningful, they must above all things be verifiable; to my mind it is just as essential that a hypothesis should also be falsifiable.

Nevertheless, as will be obvious from our survey, a specific proof or disproof, however cogent in itself, by no means suffices to secure the acceptance of a given hypothesis. Two further requirements have almost always to be fulfilled. First, the new hypothesis must not be left standing in chilly isolation: it must be fitted into the general context of current scientific opinion[2]. This may entail a drastic revision of certain concepts or conclusions formerly accepted; more often it amends the older theory by demonstrating the limits beyond which it breaks down; most frequently of all perhaps it serves to coordinate and integrate what is already known. Secondly, the hypothesis must prove itself capable of fruitful developments in fresh directions—pregnant with new problems, new inferences, new lines of observation and experiment. An ad hoc hypothesis, whose application is confined solely to the narrow field in which it was first advanced, is nearly always suspect. The history of science is replete with valid hypotheses validly established—Planck's first adumbration of the quantum theory,

[1] Thus the environmental theory, the theory of conditioned reflexes, and the various types of psychoanalytic theories, are formulated by their more enthusiastic exponents in such broad and elastic terms that there is hardly anything they cannot explain. The thoroughgoing environmentalist, for instance, wholly ignores or summarily dismisses the possibility of innate or hereditary factors, and so can argue that, since all human actions are influenced by environment, environmental influences will account for each of them. Similarly the concept of repressed complexes is framed in such a way that almost every type of behaviour can be interpreted by means of it: if Tommy seems shy and over-anxious, his inferiority feelings have momentarily overcome the repression; if he is boastful and over-assertive, the repression has resulted in over-compensation. In fact, the pseudo-scientist plainly believes that the hall mark of a perfect scientific theory is that it shall be so constructed as to baffle any attempt to negative it.

[2] Psychology again provides a salient instance: the hypothesis of extra-sensory perception has been amply confirmed by the researches of Rhine, Soal, and their followers; but, since it has hitherto defied every effort to incorporate it intelligibly into the general background of psychological knowledge, few psychologists are as yet prepared to accept it. Often, as past experience shows, the ultimate acceptance of such novel hypotheses depends, not on some one spectacular proof, but rather on a gradual transformation of the whole climate of opinion. It is this change in the general background of knowledge rather than any direct empirical disproof which, as the history of science plainly shows, has led to the final rejection of pseudo-sciences, like astrology, alchemy, oneiromancy, and the belief in witchcraft.

Mendel's particulate theory of inheritance, Freud's theory of unconscious processes as exemplified in dreams—each for many years neglected or denied, until their novel characteristics were shown to be highly suggestive when applied in entirely fresh directions.

These two requirements play into each other's hands. The combination of the new hypothesis with old facts and theories provides the joint premisses for the novel and unexpected applications; and the result is the subsumption of two or more relatively specialized hypotheses within a wider and more comprehensive superhypothesis, which serves as a general 'law' from which the former can be deduced. The process is repeated: and the final outcome is a hierarchy of logically related hypotheses, organized into one coherent deductive scheme[1].

5. The view that the hypotheses which science seeks to establish are, or ought to be, essentially 'causal hypotheses' calls for only a passing notice here. In this paper I am concerned primarily with logical structure, not with concepts or contents. So far as theoretical science is concerned, a survey, far more complete than our own, would show quite clearly that, in the strict sense of the term, the postulate of causation is never required and seldom used[2]. On the other hand, in applied or practical science it provides a helpful and convenient way of stating certain relevant facts and regularities. Even so, however, we are by no means bound to assume (as Professor Smith would have us do) that "the causal determinants [of an effect] must be spatially and temporally continuous with it". As we have already seen, Newton and many others explicitly rejected such a doctrine. In psychology certainly it has proved a hindrance quite as often as a help; and we should, I think, feel perfectly free to surrender it if we wish. However, I have discussed the causal postulate more fully elsewhere in this issue, and therefore need enlarge no further on it here[3].

[1] Notable instances are the way in which Einstein's doctrine of relativity and Schrödinger's development of wave mechanics in quantum theory allow the classical theory of dynamics to retain its validity as an approximate statement—the simpler equations appearing as limiting cases of the more complex, while the more elaborate versions clarify many of the puzzles and paradoxes which the simpler doctrines left unsolved.

[2] Professor Smith maintains that, different as they are, all the six psychological systems that he has examined assume the 'causal hypothesis'. But the only British writer whose 'system' he has examined, namely, McDougall, definitely rejects it—in this respect following the line taken by Ward, James, and Stout. McDougall again and again insists that the teleological doctrine that he himself upholds is quite incompatible with such a postulate (e.g. *Modern Materialism*, pp. 5f., 184–218).

[3] Pp. 84–5 below. No doubt, many physicists, even today, would agree with Newton's critics in holding that continuity was "an appearance to be saved" whenever possible: the concept of a gravitational 'field' and the recent theory of the 'graviton' have been put forward largely on this ground. At the same time, in view of the discontinuous nature of subatomic phenomena, it is difficult to see what precisely spatio-temporal continuity would mean at that level, which is after all commonly regarded as the most fundamental. Perhaps the most plausible explanation of gravitational forces is the suggestion that energy may be transmitted by the interchange of certain particles (e.g. mesons), much as electrical forces, which also act at a distance, may be explained by the transmission of energy by photons.

VI. Conclusions

The Structural Scheme

By way of summary let me now outline the general plan on which, if we may trust the foregoing analysis, an efficient scientific system is commonly constructed. In recapitulating the essential points, I shall try to set out the conclusions in such a way that they may incidentally offer some practical guidance to those who are planning researches of their own.

The earliest writers on methodology—those who first attempted to hammer out a cogent logical technique for the natural sciences—rightly noted that any successful enquiry necessarily involves two principal phases, which they termed *analysis* and *synthesis*. The classical statement is that appended to prop. 5 of *Euclid*, bk. XIII (a book seldom read by the modern student)[1]. Analysis consists "in *assuming* the hypothetical conclusion to be true and deducing the necessary consequences"; if these consequences turn out to be false (as in a *reductio ad absurdum*), then the assumption must have been false; if they are true, the *synthesis* will consist essentially in "reversing the analytic process". Thus, when the intermediate propositions are all convertible, the crucial steps have merely to be re-stated in the opposite order; when they are not, a more circuitous route must be patiently worked out. In adapting this general plan for the purposes of the empirical sciences, as distinct from the deductive, later commentators interpolated a third stage, namely, *heuresis* (*via inventionis*, 'the way of discovery'), in which the consequences are tested by *ad hoc* observations or experiments; and several of the schoolmen—notably Grosseteste, Ockham, and Roger Bacon—discussed in some detail the appropriate precautions to take when the mediating propositions are not convertible.

It will be seen that the general structure which we have encountered in our foregoing review confirms this basic scheme. It involves a double movement (*duplex via*): "and the road", as Bacon says, "lies not on a level, but ascends and descends". There is first the road up the pyramidal hill (*via resolutionis*); then a pause for a survey (*circumpectis*) at the top; and lastly the more devious pathway downwards (*via compositionis*). In actual practice, work may be going on at all the different stages at once; and in the final exposition brevity will doubtless demand the omission of the obvious, and a good many short cuts[2].

[1] The distinction cited in the text is traditionally attributed to Plato and the definition to his pupil, the astronomer Eudoxus. The rather inappropriate Latin translations, *resolutio* and *compositio*, were introduced by Chalcidius (4th cent.) in his commentary on Plato's *Timaeus*. The doctrine, as formulated above, formed the basis of the discussions on scientific method which started afresh at Oxford with Grosseteste and his school; it is accepted by Galileo, who continually refers to *il metodo resolutivo* (or *ex suppositione*) and *il metodo compositivo*: (cf. the famous passage in which he points out that, to save repetition, so many writers tend to give only one stage of the argument, and leave the reader to infer the other: "Thus Pythagoras, long before he found the demonstration for which he sacrificed the ox, must have reached and verified the conclusion there proved": *Due Massimi Sistemi, Giorn. Ia*, 75–6).

[2] The common mistake of the research student is to overlook the need for the preliminary 'analysis'. Mill's influence induced modern writers to rename the two stages 'deduction' and 'induction'. But these names are misleading: both types of inference may appear in both phases, as is shown by the fact that some writers describe the first as inductive and the second as deductive, instead of vice versa.

Here, however, we are concerned solely with the logical requirements, not with the expository; and this leads us to distinguish a number of substages within each main stage.

A. The Analysis

1. A scientific enquiry starts, not with a collection of facts to be generalized into a hypothesis, nor yet with a 'hypothesis' to be corroborated by the facts, but with a *problem.* In every research the most difficult task—and one too often skipped—is the preliminary formulation of the exact question to be answered. I suggest that, as a rule, this can be done most effectively if the question is expressed in disjunctive form, explicitly specifying the alternative solutions.

2. In considering how to determine which solution is to be accepted, the first step will be to decide what methodological procedure is appropriate and practicable. That in turn will depend on our choice of a *logical model.* According as its constituent elements are qualitative or quantitative, discrete or continuous, scalars, vectors, matrices, or operators, so the method to be employed will obviously differ.

3. The entire discussion will involve the enunciation and the logical arrangement of a chain of *propositions.* In general there will be several distinguishable types[1].

Propositions Describing the Abstract Model

(i) First, we need to know what *postulates* are essential to describe the theoretical model that has been chosen. Save in exceptional circumstances, both model and postulates will be more or less familiar to the reader. Hence an explicit enumeration may not be needed in the published exposition; but the investigator will do well always to formulate them for himself.

(ii) Next, we need formal *definitions* for all the relevant concepts. These will be stated in terms of other concepts which may (*a*) be intelligible to the reader without further explanation, (*b*) require to be defined in their turn, or (*c*) be treated as undefined, but subsequently interpreted by 'ostensive definitions' or the like.

(iii) Thirdly, the *hypotheses* to be examined must be stated in a form that lends itself to rigorous proof or disproof by the method of enquiry selected.

[1] It is not my intention to prescribe any stereotyped procedure, fool-proof for every kind of problem. Nor do I suggest that, in casting his argument into its final shape, the investigator should formally label every stage and every statement with its logical designation, though this is a common practice in mathematical papers. He should, however, carefully note for himself what is the logical function of each crucial proposition or concept, and make it easy for the critical reader to do so if he wishes. I fully concur with Dr. Woodger when he declares that, in psychology as in other sciences, "it is highly desirable to cultivate a sense of awareness of the type and status of the statements we are using" ([54], p. 32f.: in this passage Dr. Woodger offers a simplified classification of his own: it differs somewhat from mine and from that which he previously put forward in his earlier book [53] which was based on the methodological theories of Russell and Whitehead).

If possible, they should be jointly exhaustive and mutually exclusive[1]. Dr. Richardson asks (this *Journal*, VI, p. 107) "how can anybody be certain that he has thought of *all* the possible alternatives?" He cannot. Usually he will have to be content with choosing those that have the highest antecedent probability; and for this (apart from some flash of original inspiration[2]) he must rely chiefly on a thorough knowledge of his own special subject.

(iv) The next step will be to deduce the actual *corollaries* to be tested. In the more complex sciences the hypotheses themselves will seldom be susceptible of any direct check. Hence it will be necessary to enquire first of all what observable consequences would follow if each hypothesis were true. To enable a crucial test to be carried out, each corollary (or set of corollaries) should be distinctive of its own hypothesis, and, if possible, should have a *low* antecedent probability, i.e. should state occurrences that would be highly improbable if the hypothesis were false.

(v) Finally, we shall need certain semantic rules or *interpretative principles*, explaining how the concepts mentioned in the corollaries are to be identified in the subsequent observations or experiments. These will often consist of what many writers would call the 'operational definitions'.

B. The Observations

Only after completing this preliminary analysis is the investigator ready to plan and carry out his actual observations. Unless they have been collected with the problem, the hypotheses, and the derivative corollaries all kept explicitly in view, the data obtained will be entirely useless. As we have already noted, blindly accumulating a mass of facts or measurements in the hope that an illuminating hypothesis will ultimately dawn, like a gleam of light through the gathering fog, is one of the commonest illusions of the young research worker.

Propositions Describing Concrete Results

The results of the experimental stage will, as a rule, need to be embodied in propositions of two main types.

(i) The first will consist of the *observational records*. Their purpose is to give the actual data in full—i.e. the crude and detailed facts with no inferential

[1] It is to be noted that, in many if not most of the controversies that have split psychologists into separate camps, the rival hypotheses are not in fact mutually exclusive, though their champions argue as if they were. As actually formulated, each commonly contains a modicum of truth; and the eventual solution is achieved by a kind of dialectical advance—'thesis' and 'antithesis' being reconciled through assimilation in some 'higher synthesis', which incorporates what is true, and rejects what is false (often mere over-emphasis), in each. The simplest way to ensure a clear-cut issue is to formulate one of the alternatives as the *contradictory* of the other (not as the *contrary*, since that always leaves a third possibility open): the most familiar instance of this device is that which sets up what Fisher has called a 'null hypothesis'.

[2] These are far rarer than research-students generally imagine: most suggestions that claim originality have in fact been anticipated by some earlier writer.

or interpretative element whatsoever (a counsel of perfection). They correspond to what the older logical positivists called *Protokollsätze*; these were supposed to be expressed in the 'basic language' describing 'immediate experience', and so to be beyond all doubt or question: in point of fact, they will usually be reported in an 'object language' which uses the familiar words of everyday life.

(ii) A second set of propositions will be wanted to state what, for lack of a better phrase, I will call the *observational summaries*[1]. The raw data obtained from the actual enquiry must be reduced or condensed to a succinct digest which will enable the results to be directly compared with the corollaries derived from the rival hypotheses. The problem of describing a complicated mass of data in as few terms as possible is greatly facilitated by using mathematical functions and statistical procedures. It need hardly be added that the type of statistical analysis to be used should itself be determined before the observations are planned. The process of reduction is bound to sacrifice some of the information; but the methods employed should be so chosen as to allow as little as possible of what is really relevant to escape through the sieve.

C. The Demonstration

The Form of the Final Proof

The last step of all is to weigh the accumulated evidence thus available and decide which of the alternative hypotheses are refuted by the actual results, which (if any) is confirmed, and to what degree. This requires a comparison of each observational summary with the corresponding corollary. In general outline the final proof will take the form of a disjunctive syllogism, resting on two hypothetical prosyllogisms. In the simplest case we might set it out symbolically as follows:

"(i) With the definitions we have adopted and with the postulates imposed by our logical model, there are two and only two mutually exclusive possibilities: either (*a*) X is A or (*b*) X is not A. If X is A, then Y will be M, but Z will not be N; if X is not A, then Z will be N, but Y will not be M. (ii) Now the observations have shown that Y is M and Z is not N. (iii) Hence X is A."

This simple type of argument appears again and again in the writings of Galileo, Newton, and their immediate followers, where the proof is commonly treated as a matter of ordinary syllogistic deduction. But by the time Gauss and Laplace had finished working over their predecessors' proofs afresh, it had become clear that the reference to empirical observations nearly always implicitly transformed what had been formulated as categorical certainties into statements

[1] Older writers and more recently Woodger ([54], p. 32) refer to these condensed assertions as 'generalizations' of the observations. The phrase is somewhat misleading. A 'general proposition' would have the form "*All* instances of x have the property y"; and (unless x is of such a nature that the number of instances is strictly finite) no set of observations could justify such a statement. The most that they will entitle us to say will be that "*Some* instances of x—namely, those (or a certain proportion of those) comprised in our observed sample of the total population—have the property y", or alternatively "The average value of y in this sample of x is $y \pm s_y$".

involving probabilities. Nevertheless, even after this important modification the three main steps or stages will be much the same as before.[1] We therefore have to consider three types of probability: (i) the 'prior' or antecedent probabilities of the alternative hypotheses, $P(h_i|H)$, say, where H denotes the information available before the experiment has been carried out. (ii) The 'likelihoods', $P(r|h_iH)$, that, if this or that hypothesis (h_i) were true, then the results observed (r) would ensue. (iii) The posterior probabilities, $P(h_i|rH)$, that, in view of the results actually obtained, this or that hypothesis is in fact true. By Bayes' theorem for inverse probabilities, the posterior probability for the ith hypothesis may in theory be determined as follows:

$$P(h_i|rH) = \frac{P(h_i|H) \,.\, P(r|h_iH)}{\sum_i \{P(h_i|H) \,.\, P(r|h_iH)\}}.$$

Decision

There are in practice well-known snags that impede the use of this theorem just as it stands. If we are to calculate (iii), then estimates for both (i) and (ii) must be available in quantitative form. In a few problems, e.g. in those dealing with Mendelian hypotheses, the antecedent probabilities can be specified exactly. In most cases, however, they will at best be little more than crude estimates of subjective credibility[2]. Yet even so, we may at times feel sure that, whatever the actual values may be, they must be widely different for the different

[1] The change in method merely involves a change from a two-valued logic (of the traditional type) to a multi-valued logic in which "the alternative 'true–false' is replaced by a continuous scale of truth values" (H. Reichenbach, *Theory of Probability*, pp. 388f. : for a defence of Bayes' theorem and its importance in empirical science, see *ibid.*, pp. 94f., and for the change back from a statistical generalization to an 'all-statement'—which Mr. Taylor finds so difficult, see p. 438).

[2] For an examination of the issues involved, see A. Wald, *Statistical Decision Functions* and R. A. Fisher, *Statistical Methods and Scientific Inference*. In recent controversies on decision procedures the disputants tend to assume that there is only *one* safe method, namely, their own. In my view, different procedures are appropriate to different types of problem. Thus with problems in practical psychology it is also desirable to take into account (what so many controversialists entirely overlook) the *cost* of securing the information and its *utility* when secured. The most recent treatment of the whole question is itself based on a postulational method, similar to that advocated for more general purposes here: cf. J. Marschak, *Econometrica*, 1950, XVIII, pp. 111–141, and H. Chernoff, *ibid.*, 1954, XXII, pp. 422–443.

Mr. Taylor maintains (in the current issue of *Brit. J. Psychol.*) that 'genuine laws of nature', such as are found in other sciences and the psychologist should seek, are "never statistical laws" and are "not discoverable by statistical methods". From his references to Newton and Einstein it seems clear that he means by laws of 'nature' laws of a hypothetical model interpreting 'nature'. But the validation of such 'laws' (i.e. determining how closely they agree with the data) is always a statistical problem (e.g. testing whether the discrepancy between expected and observed changes in Mercury's orbit exceeds the probable error). He defends his view by arguing that Newton's laws dealt with single systems (the sun and a planet, the earth and a falling weight), not with averages, aggregates, or frequencies. All such bodies, however, *are* aggregates of a vast number of elementary particles ; and Newton's 'laws' are approximations only, which hold good of actual bodies only when the variables are averages based on large numbers, and even then solely within certain limits (for proof see C. G. Darwin, *Proc. roy. Soc.*, A, CXVII, p. 258). Moreover, the 'laws' of the kinetic theory of gases and of Mendelian inheritance (to mention but two examples) would never have been discovered *without* statistical methods (cf. [28] and [32], esp. chaps. 6 and 18).

alternatives (in psychology the most glaring instances are the *a priori* probabilities of hypotheses affirming and denying telepathy, clairvoyance, and other paranormal processes). If no evidence whatever is available as to their relative size, the traditional treatment proceeds by assuming that they are equal. On the other hand, the likelihoods, at least in any well-planned research, can usually be determined from the data; and in that case, when no information is forthcoming about the prior possibilities, the most obvious working principle is that of 'maximum likelihood'.

But no matter what the precise situation may be, before he can decide whether any conclusion at all can validly be drawn, and if so what it is, or whether more extensive enquiries are needed, some definite principle must be formulated in advance stating how, in the problem before him, the investigator tends to adjudicate between the various possibilities. Thus, as we have seen throughout this discussion, at each successive stage down to the very last, it is only by scrupulously considering the details of his procedure before ever he embarks on his actual experiments that a scientist can hope to make any solid contribution to knowledge,

> Und was in schwankender Erscheinung schwebt,
> Befestigen mit dauernden Gedanken.

References

[1] Ashby, W. R. (1957). *An Introduction to Cybernetics.* Chapman & Hall.
[2] Ayer, A. J. (1954). *Language, Truth, and Logic.* Gollancz.
[3] Boring, E. G., Langfeld, H. S., and Weld, H. P. (1948). *Foundations of Psychology.* John Wiley.
[4] Braithwaite, R. B. (1953). *Scientific Explanation.* Cambridge University Press.
[5] Bridgman, P. W. (1927). *The Logic of Modern Physics.* New York: Macmillan.
[6] Broad, C. D. (1919). Mechanical explanation and its alternatives. *Proc. Arist. Soc.*, XIX, 86–124.
[7] Broad, C. D. (1947). *Mind and its Place in Nature.* Kegan Paul.
[8] Broad, C. D. (1923). *Scientific Thought.* Kegan Paul.
[9] Broglie, Louis de (1946). *La Physique Nouvelle et les Quanta.* Paris : Alcan.
[10] Burt, C. (1937). *The Backward Child.* University of London Press.
[11] Burt, C. (1938). The unit hierarchy and its properties. *Psychometrika*, III, 151–168.
[12] Burt, C. (1940). *The Factors of the Mind.* University of London Press.
[13] Burt, C. (1955). The evidence for the concept of intelligence. *Brit. J. educ. Psychol.*, XXV, 158–177.
[14] Burt, C. and Howard, M. (1956). The multifactorial theory of inheritance and its application to intelligence. *Brit. J. statist. Psychol.*, IX, 95–131.
[15] Carnap, R. (1926). *Physikalische Begriffsbildung.* Karlsruhe: Braun.
[16] Cassirer, E. (1922). *Das Erkenntnisproblem in der Philosophie und Naturwissenschaft der neueren Zeit.* Berlin: Cassirer.
[17] Chambers, E. G. (1943). Statistics in psychology and the limitations of the test method. *Brit J. Psychol.*, XXXIII, 189–199.
[18] Duhem, Pierre (1914). *La Théorie Physique: Son Objet, Sa Structure.* Paris: Rivière.
[19] Eddington, A. (1928). *The Nature of the Physical World.* Cambridge University Press.

[20] Eysenck, H. J. (1952). *The Scientific Study of Personality*. Routledge and Kegan Paul.
[21] Frazer, R. F., Duncan, W. J., and Collar, A. R. (1938). *Elementary Matrices and Applications to Dynamics*. Cambridge University Press.
[22] Heim, A. (1954). *The Appraisal of Intelligence*. Methuen.
[23] Hilbert, D. (1918). Axiomatisches Denken. *Math. Annalen.*, LXXXIII, 405–419.
[24] Hull, C. L. (1943). *Principles of Behaviour*. New Haven: Yale University Press.
[25] Hutten, E. H. (1950). *The Language of Modern Physics*. Allen & Unwin.
[26] Jevons, W. S. (1900). *The Principles of Science*. Macmillan.
[27] Kraft, V. (1925). *Die Grundformen der Wissenschaftlichen Methoden*. Leipzig: Barth.
[28] Lindsay, R. B. (1951). *Concepts and Methods of Theoretical Physics*. New York: Van Nostrand.
[29] Mach, Ernst (1883). *Die Mechanik in ihrer Entwicklung*. Leipzig: Barth.
[30] Maddox, H. (1957). Nature–nurture balance sheets. *Brit. J. educ. Psychol.*, XXVII, 166–175.
[31] Margenau, H. (1950). *The Nature of Physical Reality*. New York: McGraw-Hill.
[32] Miles, T. R. (1957). On defining intelligence. *Brit. J. educ. Psychol.*, XXVII, 153–165.
[33] Murphy, G. (1956). *A Historical Introduction to Modern Psychology*. Kegan Paul.
[34] Pearson, K. (1900). *The Grammar of Science*. A. & C. Black.
[35] Planck, M. (1938). *The Universe in the Light of Modern Physics*. Leipzig: Barth.
[36] Poincaré, H. (1904). *La Science de l'Hypothèse*. Paris: Alcan.
[37] Poincaré, H. (1908). *La Science et la Méthode*. Paris: Alcan.
[38] Popper, K. (1935). *Die Logik der Forschung*. Wien: Barth.
[39] Popper, K. (1958). *The Poverty of Historicism*. Routledge & Kegan Paul.
[40] Pratt, C. C. (1939). *The Logic of Modern Psychology*. Macmillan.
[41] Richmond, W. K. (1953). Education measurement: its scope and limitations. *Brit. J. Psychol.*, XLIV, 221–231.
[42] Robinson, R. (1950). *Definition*. Oxford University Press.
[43] Russell, B. (1927). *The Analysis of Matter*. Kegan Paul.
[44] Russell, B. (1948). *Human Knowledge: Its Scope and Limits*. Allen & Unwin.
[45] Ryle, G. (1949). *The Concept of Mind*. Hutchinson.
[46] Sigwart, C. (1895). *Logic*. Sonnenschein.
[47] Smith, F. V. (1951). *The Explanation of Human Behaviour*. Constable.
[48] Spearman, C. (1927). *The Abilities of Man*. Macmillan.
[49] Stout, G. F. (1909). *Analytic Psychology*. Sonnenschein.
[50] Synge, J. A., and Griffith, B. A. (1949). *The Principles of Mechanics*. New York: McGraw-Hill.
[51] Weyl, H. (1921). *Raum. Zeit. Materie*. Berlin: Springer-Verlag.
[52] Wiener, N. (1949). *Cybernetics*. New York: Wiley.
[53] Woodger, J. H. (1937). *The Axiomatic Method in Biology*. Cambridge University Press.
[54] Woodger, J. H. (1956). *Physics, Psychology, and Medicine: A Methodological Essay*. Cambridge University Press.
[55] Zangwill, O. L. (1950). *An Introduction to Modern Psychology*. Methuen.

NOTES AND CORRESPONDENCE

SECONDARY SCHOOL SELECTION

Sir,—As teachers with a varied experience of current methods of selection, we feel that the reviewers of the British Psychological Society's report on *Secondary School Selection* have approached the problem with a decidedly narrow outlook, and have so failed to do justice to the inquiry on at least three important points. Their own conclusions seem, at least in our opinion, to illustrate the fallacies to which a purely statistical psychology is liable when left unchecked by a more humane and (as they would doubtless hold) a non-statistical and therefore unscientific, standard of judgement.

1. To begin with, they apparently cling to the out-of-date Galtonian doctrine that the population can be divided for educational purposes into two unequal sections—an aristocratic élite and all the rest: higher education—at the grammar school and the University—is to be the privilege of the élite alone. Professor Burt, who shares the same standpoint, treats the whole problem as a question of numbers, and has even computed, by processes that are left unexplained, an actual number for the élite—viz., $7\frac{1}{2}$ per cent in each age group, who are apparently to be progressively sorted out and segregated. The contributors to the Society's report, on the other hand, rightly repudiate the notion of an intellectual aristocracy and adopt an attitude that is frankly democratic. " If ", they say, " children are to learn the ways of democracy in which everybody counts, there is a defensible case for them to gain experience by working with children of many grades of ability ". Accordingly, " in areas where it is possible to make a clean break with the traditional organization ", they urge the policy of ' comprehensive ' schools as offering a " new and promising approach which avoids difficulties of streaming and segregation ". This, we believe, is as sound in theory as it is desirable in practice.

The aristocratic theory is based on the assumption of large innate individual differences; the democratic theory on the assumption that (save in rare pathological cases) innate differences between one individual and another are too slight to be of practical consequence, and that the superficial differences between social classes are artificially produced. In any case it is plainly the preferable assumption for a democratic community to adopt. In such a community " no individual has a right to more of this world's benefits than any other "; and that, we hold, should apply to education as to all else.

In America we have before our eyes an outstanding example of the way in which a genuinely democratic scheme of education can serve the needs of the nation. There, though the vogue of mental testing and statistical analysis is quite as conspicuous as in Britain, it has not been allowed to distort educational aims by an eleven-plus examination or by insisting on a privileged type of education for the privileged few. There, as the writers of the Society's report observe, " the prestige of the grammar-school type of education is not something firmly entrenched ", as it is in the mother-country.

2. No doubt, so long as a large proportion of the child population continues to suffer from the handicaps of bad health, bad housing, and an inadequate cultural background, many pupils will turn out to be less fit for a grammar school education than the rest. If, as the Act requires, some kind of selection process is for the present rendered almost inevitable, then the merits of that procedure must be judged by its validity or accuracy; and, contrary to general opinion, this, we believe, does not justify selection procedures such as those which psychologists have hitherto favoured. The reviewers, however, apparently hold that the proper policy is to judge procedures in terms of actual cost. And the Report is criticized because it relies on a straightforward validity coefficient instead of some ancient and highly sophisticated ' statistical decision procedure ', which was suggested some thirty years ago and which involves an extremely speculative formula for estimating ' net gain ' by balancing expected profit and cost.

The formula the reviewers prefer is based on the Bayes–Laplace–Pearson principles developed by early theorists to answer the questions posed by people whose chief interests lay in gambling, dicing, and betting; and indeed Professor Burt seems perfectly satisfied to treat the whole matter as a vast gamble. But the principles themselves, together with Bayes' postulate and all the rest, have surely been disposed of by Sir Ronald Fisher. " It is ", he writes, " important that the scientific worker introduces no cost functions for faulty decision . . . The idea that responsibility can be delegated to a computer programmed with ' decision functions ' belongs to a fantasy of circles remote from scientific research " (*Statistical Methods and Scientific Inference*, pp. 100f.). If, therefore, the writers of the Report ignored all such proposals, that surely shows that they realized that a validity procedure offered a far sounder approach than any so-called ' decision ' or ' acceptance ' procedure.

3. The reviewers criticize the Report on the ground of the inconsistencies they discern, or think they discern, in chapters written by different authors. But the apparent inconsistencies really stem from the much broader standpoint which the Working Party has adopted. They vanish altogether when we remember that the needs of a transitional stage in educational reorganization must necessarily differ from those which can be fulfilled only when reorganization has been more or less completed. In any case, the reviewers' own arguments are still more inconsistent. They accept the definition of intelligence as inborn general ability; yet they admit that the intelligence tests used for selection at 11 plus do not furnish accurate measures of inborn ability, and result in the misplacement of at least 10 per cent. On their own premisses, therefore, the arguments for discarding such tests would seem to be far stronger than those for retaining them.

W. G. Hopkins.
A. K. Marshall.
C. F. Johnson.

SECONDARY SCHOOL SELECTION: REPLY

We fully agree with the basic principle expressed by the three writers of the foregoing letter, namely that ' statistical psychology ' cannot of itself determine educational policy: the views of ' teachers with varied experience ' should undoubtedly be accorded their full weight. The trouble is, however, that the views of teachers differ so widely that some objective means of deciding between their theories and their predictions is essential. When one headmaster tells us that " even from schools in good areas, less than 3 per cent of the pupils appear to be capable of the work for even a low class in an honours degree ", while another maintains that the proportion is nearer 15 per cent, only a statistical inquiry can decide which is nearer the truth. Mr. Hopkins and his colleagues protest against the treatment of such problems as a " question of numbers ". But if a University is to provide more places, or an education authority is to build more grammar schools, the first question to be asked is, and must be, " how many ? ".

However, what the writers really object to would seem to be the procedures " which psychologists have hitherto favoured " and the alleged assumptions on which those procedures rest. Accordingly, may we take their three main points in turn ?

1. They begin by criticizing the attempts made by Burt and others whom we quoted to " compute an actual number for the élite by processes that are left unexplained ". By ' the élite ' presumably the letter merely means (in the words of the French dictionary) ' the part chosen '. The figures quoted for the proportion of the population to be chosen were based, not (as the letter seems to suppose) on some ' Galtonian doctrine ', but (as was explicitly stated) on the recommendations of the Scientific Manpower Committee of the Advisory Council on Scientific Policy. As endorsed by the Chancellor's statement of

20 February, 1958, these recommendations contemplate a rise in the total population of the Universities from its present figure of 95,000 to one of 124,000 by about 1965, with a possible increase to 136,000 by 1970.

The calculations leading to the estimate of $7\frac{1}{2}$ per cent are quite simple. According to the official figures, out of about 645,000 who form the 17-year age-group, nearly 10 per cent (about 63,000) are at present in the sixth forms or their equivalent. By 1965 the number who are then 17 will be increased to over 930,000; but by 1969 it will have dropped to 730,000 after which there will be once again a slow and gradual increase. Since the birthrate is still rising, the annual increase will presumably continue for some time. To judge from present trends, therefore, we may anticipate that by 1968 the number in the sixth forms of the grammar schools will be about 110,000—roughly 15 per cent of the age-group. By then about half that number—say 55,000 or about 7·5 per cent of the age group—should be proceeding annually to the Universities, Technical Colleges, and Teachers' Training Colleges. The figure of $7\frac{1}{2}$ per cent therefore that we cited from Burt arises out of a specific practical requirement, not from any *a priori* theory of an ' aristocratic élite '.

The programme proposed would obviously entail a marked advance in the provision for higher education in this country; and, if an adequate return is to be secured for this outlay, it is plainly essential that those selected should consist of the ablest individuals of each generation. A borderline of +1 S.D. cuts off just over 15 per cent of a normal group. Hence on the conventional scale the line of demarcation for this proportion would be about 115 I.Q. Many of the brighter pupils, however, will no doubt be diverted before the age of 17 into industry and commerce. Consequently, if teaching standards are not to be lowered the processes of selection will require to be, not relaxed, but if anything made more rigorous and effective.

Our critics, however, maintain that education in a democratic community should be based on the assumption that " innate differences between one individual and another are too slight to be of practical importance ", i.e. on an assumption of mental equality. This, they maintain, is the ' preferable assumption '. But as a basis for action an assumption is ' preferable ' only if there are reasonable grounds for thinking that it is probably true. In our view there is convincing evidence for the opposite doctrine, namely that the range of individual differences in innate mental ability is quite as large as it is in height or in the shape of the face. If so, it becomes essential for those who seek to lay down an educational policy to take such differences explicitly into account. Many of us no doubt would agree that a world in which all men were born equal was to be ' preferred ', just as we might ' prefer ' a world in which there was no such thing as physical weakness or bodily disease. But preference is no substitute for a realistic study of the actual facts.

The interpretation of the ' democratic principle ' which the writers propose seems to us inimical to the best interests of the community as a whole. " No individual ", they say, " has a right to more of this world's benefits than any other; and that . . . should apply to education as to all else. " Now it can hardly be denied that at least half the population could not possibly profit by a course for an honours degree; and, if we adopt the principles advocated in the letter, this would seem to mean that even those who are capable of profiting by such a course have no ' right ' to this particular ' benefit '! However, here we appear to be travelling beyond the problems of psychology as such.

The letter goes on to cite the experience of American education as demonstrating the practical benefits of the policy which the writers advocate. We question whether American educationists themselves would endorse their views. The average child certainly gets a better chance in America than in England; but the same cannot be said of pupils with higher ability. This neglect of the ' gifted child ' has been emphasized for many years past by some of the most eminent psychologists in the United States—notably Whipple, Terman, and their collaborators; and recent international events have brought the problem much to the fore. During his work as Chief Psychologist in the U.S. War Department, Walter Bingham became so much impressed by the " national neglect of its ablest members " that he bequeathed a large fund for publicizing the subject and for encouraging statistical

research on the whole complex problem: " our budding geniuses ", he wrote, " are the most valuable assets that the nation possesses, and to waste them is an act of the gravest national folly ".

Thus, strangely enough, while Mr. Hopkins and his colleagues are asking us to copy the pattern of American education, American educationists are insisting that America has much to learn from the way Britain has for a century or more sought out and aided the ' gifted child ' by a system of scholarships and free places. " At present ", writes Professor Heck, " our schools are organized for the child of average ability; even the handicapped are cared for adequately; it is the gifted youth who has become the forgotten child " (*The Education of Exceptional Children*, p. 424). Still more recently Drs. DeHaan and Havighurst have issued a review of the whole question together with results of their own enquiries in a book on *Educating Gifted Children*. And indeed during the last six months the sensational achievements of Russian technologists have precipitated quite a spate of self-critical pronouncements.

Curiously enough, as Professor Lauwerys has recently pointed out, it is Russia and the so-called communist countries that claim to have based their educational policy on a " theory of complete intellectual equality "; but, as a Russian educationist has recently remarked, " ' equality ' is to be understood in a civic, not a mathematical sense; we are primarily citizens and comrades, not mathematicians : . . . we are realists as well as idealists ". Moreover, as one of us is inclined to believe from somewhat limited first-hand observation, although the Russian ideal is one of cooperation rather than favouritism or competition, there is in practice a rigorous and progressive selection, which somehow seems (at least on the surface) to have succeeded in avoiding inferiority complexes.

To advocate " a policy of comprehensive schools " for the country as a whole seems to us quite unrealistic. Probably in certain types of area (e.g. large industrial areas), they are more economical and more effective; in others they would seem to be impracticable. But their merits are administrative and financial rather than social or political; and most of the unbiased advocates who have supported them have commonly laid chief stress on the way in which classification and reclassification of children according to abilities and aptitudes would thereby be facilitated, not on the way in which they might be avoided. In any case, the mere fact that " the sons of professional families in X-town attend the same school as sons of casual labourers ", when in point of fact they are not likely to get allocated to the same classrooms and in so huge a school are not likely to meet out of the classrooms, can scarcely be expected to affect the ' democratic attitude ' of either group in after life.

2. The next issue raised by the writers of the letter is our suggestion that the Society's Working Party, instead of assessing the merits of intelligence tests and other methods of selection merely in terms of validity coefficients, should have adopted some statistical ' decision procedure ', similar to that used by Burt in his work for the London County Council. They appear to believe that the Working Party silently rejected the idea, partly because it is ancient and therefore out-of-date, and partly because of the complexity of the formulae required.

The underlying principle can hardly be regarded as out of date, since the most recent and the most thorough discussion of this problem (emanating as it happens from America) expressly recommends the use of such a procedure in place of the simple validity coefficient. The monograph in question (*Psychological Tests and Personnel Decisions* by Cronbach and Gleser) was, in point of fact, reviewed in an earlier issue of this journal, and is now given a full critical notice in the *British Journal of Educational Psychology* from the pen of the editor of the Working Party's report. The statistical methods and conclusions put forward by the American authors, says Professor Vernon, " deserve to be pondered by all who are concerned with educational and vocational assessment, selection, and guidance ". Along such lines, he believes, " it is possible to work out mathematically the most efficient decision-making policy "; and he suggests that " in selection at 11 plus we might arrive at some such formula " as that which he writes out in detail. It consists in calculating ' the gain by the production of a highly educated *élite* selected by the tests ': this is equated to

(*a*) the advantages due to the superior validity of the tests as compared with that of the existing procedures (duly weighted) *minus* (*b*) the disadvantages accruing from such items as ' cost of test-construction and administration ' and other somewhat imponderable factors—in short (to quote the phrase to which our critics so strongly object) " by balancing expected profit and loss ". Except that the terms to be added, subtracted, and multiplied are expressed in words rather than in algebraic symbols, Professor Vernon's equation, it will be seen, is very much the same as that which we ourselves suggested, and is certainly quite as complex.

He goes on to state that the book reviewed " represents a long-overdue revolution in testing policy ". The practical implementation of such a policy is certainly overdue. But the details of the policy and its importance have repeatedly been stressed by the earlier writers, both in this country and in America, whom we ourselves quoted. We may add that such methods have frequently been used with considerable success for analogous problems in industry.

Mr. Hopkins and his colleagues cite Professor Fisher's criticisms of the use made by certain scientific investigators of ' decision ' and ' acceptance ' procedures. But they fail to observe the distinction which Fisher clearly lays down between scientific research and practical applications (*loc. cit.*, p. 76). In the latter, particularly with any ' large concern ', it is, he says, essential to " keep low *both* the frequency with which satisfactory lots may be wrongly rejected *and* the cost of testing " (our italics); and he emphasizes that he himself does not for a moment " disdain an artifice of proved value in commerce and technology ". What he criticizes are the proposals of Wald and Neyman to transfer such methods to the theoretical problems of pure science. Mr. Hopkins quotes the first and (irrelevant) half of Fisher's final conclusion. Fisher's own words are: " it is as important that the scientific worker introduces no cost functions . . . , as it is reasonable and often necessary to do with an acceptance procedure "; i.e. the consideration of cost has no bearing on the inferences drawn in *theoretical* science, but it may have an important bearing on the procedures adopted in *practical* applications[1].

3. Thirdly, our critics point out what they consider an inconsistency in our argument: we accept the definition of intelligence as inborn general ability (they say), and yet we hold that the so-called ' intelligence tests ' used for selection do not furnish accurate measures of inborn ability. But there is no real incompatibility between the two opinions. It is unquestionably desirable to provide an advanced type of education for children of high innate ability when they reach an appropriate age. But the estimate of the child's innate ability, which the Working Party considered should be made ' at the *end* of the primary stage ', ought (as we argued in our review) to be undertaken at the *beginning*, and, when necessary, checked and corrected during the intervening years. Once the children have reached the age at which the selection has to be carried out, the aim of the examination and other procedures is no longer to measure innate ability in and for itself, but rather to assess the all-round suitability of each pupil as he actually is at that stage. Probably the confusion is due to an unfortunate ambiguity in the word ' intelligence '. As we said in our review, it would be much better if, in the selection examination, what are commonly called ' intelligence tests ' were re-named ' general classification tests '.

A. G. Houghton.
D. R. Morris.

[1] It is perhaps scarcely necessary to take up the point about " early theorists whose chief interest lay in gambling, dicing, and betting ". Here as elsewhere the writers of the letter are seeking (so at least it would appear) to stir up an unfair prejudice by the phrases they employ. The fact that the Chevalier de Méré was a rather discreditable gambler does not of itself bring discredit on the answers given by Pascal. Respectable ecclesiastics, as well as gamesters, have appreciated the " fact that, in most practical problems, statistical reasoning rather than the logic of the idealist is likely to yield the safest answer ". " Probability ", said Bishop Butler, " is the guide of life "; and Dean Inge observed that " nearly all public decisions are of the nature of a gamble ".

QUANTUM THEORY AND THE PRINCIPLE OF INDETERMINACY

To the Editor, British Journal of Statistical Psychology

SIR.—May I lodge a mild protest against the way in which psychology, particularly psychology as interpreted by the exponents of factorial procedures, has tended of late to depart more and more from the experimental study of determinate causes, and to content itself with mere statistical studies yielding nothing but vague probabilities ? In the recent papers on ' stochastic processes ', for example (*Brit. J. statist. Psychol.*, IX, 49–62 and 87–94), all the investigators appear to take it for granted that " in psychology deterministic laws must be replaced by probabilistic laws, . . . laws expressed in terms of probabilities, not certainties ". And Professor Burt would have us believe that " explanation in terms of causal agencies is a legacy of nineteenth century science—a mode of approach which should be regarded as out of date in mental science just as it is in physical science " (*Factors of the Mind*, p. 218).

Such statements, however, seem flatly to contradict the views of the vast majority of present day psychologists. Hebb, for example, assures us that ' modern psychology ' takes ' complete causation ' for granted (*The Organization of Behaviour*, p. xiii); and Professor F. V. Smith, after reviewing the leading psychological ' systems ' of recent years, points out that all of them " accept the position that human behaviour is *determined* " (*The Explanation of Human Behaviour*, p. 257).

The reaction against causality seems ultimately to have been inspired, not by psychologists, but by physicists like Jeans and Eddington. Their chief ground for extending it to psychological problems was the belief that " the principle of indeterminacy would go far to explain the paradoxes of introspection and to remove the difficulties connected with free will ". The philosophers seem at first to have followed suit. Professor Broad, for instance, almost at once endorsed the Eddingtonian inferences: in discussing what he proposed to call the ' non-causal characteristics ' of mind and matter, he stated that, " if the Quantum Theory be correct, we are probably witnessing a breakdown of the accepted assumption "—the assumption, namely, that the ' minute structure ' of the brain, as of other physical substances, is subject to the same laws as the ' gross structures ' which we actually observe (*The Mind and its Place in Nature*, 1925, p. 435).

But all this was over thirty years ago. In the meantime Einstein himself has produced a cogent argument which, so it would seem, has removed the indeterminacy from physics; and Professor Stebbing has dealt still more drastically with the alleged philosophical inferences. In psychology, so far as I am aware, the only thorough discussion of the question is that of Professor J. F. Brown (*Psychology and the Social Order*, Appendix A). " In every modern discussion of methodology ", he writes, " the problem of determinism and indeterminism, of dynamical analysis and statistical, has been stressed. . . . Several reputable physicists and biologists have interpreted the indeterminacy relation as indicating that the biological sciences must be content with statistical analysis, expressing predictions in terms of probability only; and on these grounds some psychologists have questioned the attempt at causality analysis in psychology. But the argument is basically unsound: it confuses the validity of a scientific method with the probability of the predictions based on it. . . . The development of quantum mechanics affords no reason for an attitude of defeatism in psychology; and in so far as we have to accept a certain metaphysical bias, we must believe metaphysically in determinism. "

Surely those statistical psychologists who still claim that the indeterminacy relation releases the psychologist from his belief in causation, should set forth their arguments in clearer detail, and at least explain (if they can) where physicists like Einstein and psychologists like Professor Brown and Professor Smith are really at fault. Professor Burt

defends his position by saying that " indeterminacy does not mean indeterminism ". But this seems little more than a quibble over suffixes.

Science throughout the past has been based on a thoroughgoing determinism; and any proposal to abandon determinism would surely be revolutionary. If in actual practice our predictions can at times state no more than the probable effects, that is due solely to human ignorance—an ignorance which the progress of knowledge will remedy—and not to the nature of the phenomena themselves. For instance, when I say that, among the first thirteen cards dealt from a well shuffled pack, I shall ' probably ' find one ace, the ' indeterminacy ' arises merely from the fact that, being neither a conjuror nor a clairvoyant, I do not know the exact order of the cards. There is, underlying such probabilities, nothing *intrinsically* indeterminate, as the arguments about stochastic processes would appear to assume. Nor can I see that anything is gained by recommending us to model our theories on an Eddingtonian hypothesis which seems to assume that the fundamental type of entity in the universe is at once a particle and a wave.

Admittedly the doctrine of strict scientific determinism is in sharp conflict with our seeming sense of volitional responsibility. But, since we are forced to choose between the two, we must surely acknowledge the superior claims of a logical determinism, and give up the dogma and the illusion of a free and arbitrary will. It is difficult to avoid the suspicion that at bottom the doctrine of indeterminism rests simply on a sentimental preference for a vague irrationality, such as would allow room for personal desire or caprice. But, if psychology is to remain a science, it must, like other sciences, preserve an unswerving belief in universal law and order.

M. L. Clarke.

Sir.—At the Reading Symposium on Factor Analysis Professor Burt castigated certain factorists for " relying on merely verbal argument to prove or disprove what are essentially mathematical theorems ". " The problems of factor analysis in psychology ", he writes, " *are seen* " (my italics) " to be precisely analogous to the reduction of a matrix to canonical form and the determination of the ' Eigen-values ' and the ' Eigen-functions ' describing the ' state ' of a physical system in quantum theory " (*Psychometrika*, III, 1938, p. 156f.). Indeed, he and Slater are prepared to follow the physicists still further by proposing the introduction of ' complex ' and ' imaginary ' quantities—a suggestion which would surely defy any attempt at operational verification by means of actual experiments (cf. Burt, *loc. cit.*, p. 151, and *Brit. J. statist. Psychol.*, IV, pp. 18f., and Slater, *ibid.*, pp. 9–17, and VI, p. 101).

To support all this Burt himself seems to rely merely on verbal arguments. Occasionally he refers to ' recent developments in physics '. But after all few students of psychology are familiar with the details of modern physics; and allusions to ' Eigen-values ', ' Eigen-functions ', ' states ' (in inverted commas), and ' waves of probability ', merely leave them mystified. Hence a paper summarizing what in the theories of quantum physics is really relevant to the problems and procedures of the psychologist, and including some clearer indication of what the ' mathematical arguments ' are, would be welcomed by many readers of your journal. Most of us are willing to believe that the time has come for the results of such theories to be incorporated in the work of the psychologist, if only we were given more convincing explanations. What we really want to know is this: is the proposal to introduce what may be called quantum mathematics based merely on vague analogies, or is it put forward because the neurological processes underlying human behaviour are considered to be themselves of a quantum type ?

L. F. Jackson.

QUANTUM THEORY AND THE PRINCIPLE OF INDETERMINACY :

The following note attempts to answer the various criticisms recently urged by Mr. Clarke, Mr. Jackson, and others (*sup.* pp. 75–6) against the views put forward in the papers they have quoted.

Mr. Clarke's main argument seems to rest on the notion that, apart from " a quibble over suffixes ", ' indeterminism ' and ' indeterminacy ' are synonymous. But in the discussions to which he refers each word was used in a technical sense. For the psychologist, ever since James' celebrated pronouncement, the word ' determinism ' has meant " the fatalistic doctrine that all actions, including what we feel to be voluntary actions based on free choice, are inexorably predetermined by the laws of cause and effect. . . . On this assumption, possibilities that fail to get realized are sheer illusions; they never were possibilities at all: in short, we really have no choice " (*The Dilemma of Determinism*, 1884). Indeterminism implies the opposite of all this (cf. Drever's *Dictionary of Psychology*). But when the physicist speaks of ' the principle of indeterminacy ' he is referring to something quite different. ' Indeterminacy ' is used as a translation of Heisenberg's word ' Unbestimmtheit '; and in the paper that first clearly established the principle (*Zeitschr. f. Physik*, XLIII, 1927, pp. 172f.), the ' determination ' referred to was determination by measurement, not by causation. Accordingly, to avoid what is a very natural confusion, it might perhaps be preferable to call Heisenberg's theorem the principle of *indefiniteness* or (as many do) ' the *uncertainty* relation '.

However, in view of the frequent references to Heisenberg's principle in recent psychological discussions, and the various ways in which it is misunderstood, it certainly seems desirable to ask what precisely his theorem maintains and how it is deduced, even at the risk of digressing into the field of quantum mechanics. This may at the same time serve to satisfy Mr. Jackson's suggestion that some account of " the details of quantum theory "—at least so far as it is relevant to the procedures and problems of the psychologist—should be made available for the ordinary student. Certainly, there are several good reasons why psychologists should give at least a passing glance at recent developments in this field. First, the results of the theory present us with a conception of the physical universe that is very different from the materialistic model of the nineteenth century to which both neurologists and behaviourists still commonly adhere; secondly, as Bohr, Schrödinger, and other quantum physicists have pointed out, the gene and the neural synapse are sufficiently small to make it probable that their essential functions ought to be conceived in terms of the quantum theory rather than the classical theory; thirdly, the study of quantum phenomena has much of the novelty, complexity, and paradoxical character that are encountered in the study of psychological phenomena[1]; and, though neither psychologists nor physicists appear to have recognized it, the mathematical methods employed in both branches of science display many striking analogies.

Mr. Jackson inquires more especially what I had in mind in speaking of ' recent developments ' in this field. I was in fact referring to what is still sometimes called the ' new ' quantum theory, though it first appeared about 1925 and has undergone surprising developments since. Its origins were curiously cosmopolitan; and in its present form it is a synthesis of three distinct lines of approach, associated with the names of Heisenberg (a German), Schrödinger (a Viennese), and Dirac (an Englishman of Swiss extraction).

The starting point was a vast array of experimental data on intra-atomic processes, which had been gathered by physicists in almost every country, and which failed to agree with the earlier hypothesis elaborated by Rutherford and Bohr. They had pictured the atom as a miniature planetary system with the electrons whirling in Keplerian orbits

[1] For a summary of the more striking analogies see D. Bohm, *Quantum Theory* ch. 8.

around a nuclear ' sun '. The basis for the ' new ' theory was an equation deduced a century ago by Sir William Hamilton and derived from the principle of ' least action '. This sought to reconcile two apparently incompatible hypotheses about light—that it is corpuscular (Newton and later Einstein) and that it is undulatory (Huyghens and later Maxwell). Prince Louis de Broglie proposed a similar dualistic theory for matter, which exhibits much the same Jekyll and Hyde behaviour as light.

De Broglie's novel suggestion was almost immediately confirmed by the experiments of Davisson and Germer in America and G. P. Thomson (son of J. J. Thomson) in this country. As is well known, in the case of light the phenomena characteristic of waves (diffraction and interference patterns) are manifested only when it encounters obstacles not much larger than the length of the waves themselves; otherwise light appears to be propagated in a straight line like a particle. In much the same way it was found that, when a beam of electrons is reflected from the surface of a crystal of nickel or passed through an extremely thin layer of gold foil, the minute lattice of atoms formed by these metallic obstacles scatters the beam at varying angles, and so produces a fringe of rings, very like that which is observed with ordinary light or with X-rays. The American experimenters used electrons of slow velocity; Thomson, electrons of high velocity. In every case the wave-lengths calculated from the resulting diffraction-patterns agreed with those predicted by de Broglie's theory.

Mr. Clarke and others naturally object to a theory which " seems to assume that one and the same material entity is at once a particle and a wave ". The simplest answer is that since the material particles with which we are familiar in ordinary life are made up of protons and electrons, protons and electrons cannot themselves be ' particles ' *in the same sense*. But the modern physicist no longer claims to describe what such a thing ' is '; all he seeks to do is to describe how it behaves. He has therefore deliberately abandoned the attempt to invent a visible or mechanical model for explaining atomic *structure* and is content to devise mathematical models for predicting atomic *events*.

1. Of the various mathematical models suggested, the one that comes closest to those used in statistical psychology is the matrix model[1] developed by Heisenberg, Born, and Jordan. Essentially their procedure consists in taking certain general equations from classical mechanics, and substituting matrices for the simple algebraic quantities. The system of ' laws ' derived in this way proves to be in close agreement with the detailed results obtained from observations on atomic spectra and the like.

One of the most remarkable consequences thus deduced may be stated as follows. Let q be a generalized coordinate and p the corresponding generalized momentum, as in the Hamilton-Jacobi equation; but, instead of scalars, let q and p represent matrices in which the elements stand for observable properties only. Now, with matrices (as every factorist knows) multiplication is not commutative. Thus $pq \neq qp$; and it may be shown that the difference $pq - qp = (h/2\pi i)I$, where $i = \sqrt{-1}$, I denotes the unit matrix, and h a fundamental constant of nature—Planck's quantum of action. It follows that we can never predict or ' determine ' both positions and the corresponding momenta with absolute exactitude. Thus, let δp and δq denote the amount of ' uncertainty ' or ' indeterminacy ' in the two measurements: then their product $\delta p \,.\, \delta q \geqq \frac{1}{2}(h/2\pi)$: in other words, the product of the two standard deviations is of the order of one quantum of action. There are several

[1] For the psychologist who wishes to study the mathematical methods perhaps the best introductory manual is Professor Temple's little book on *The General Principles of Quantum Theory* (Methuen Monographs, 1934): this brings out very clearly the relations between matrices and operators. For a non-mathematical approach an excellent introduction is Heisenberg's *Die Physik der Atomkerne* (or the English translation, *Nuclear Physics*, Methuen, 1953; see pp. 30f. for " The relation called ' the uncertainty principle ' "): it may be noted that Heisenberg here speaks with full approval of Schrödinger's methods (pp. 35f.). For the general implications see Heisenberg, *The Physicist's Conception of Nature* (1958) and D. Bohm, *Quantum Theory* (1952) esp. ch. 8.

other complementary ' canonically conjugate ' pairs of ' observables ' which do not commute, and for which therefore the same conclusion holds good—e.g., energy and time.

This then is Heisenberg's ' relation of uncertainty '. It implies that, at any rate in principle, all that the physicist's formulae can predict is the *probability* of finding at the point or the time specified a particle possessing this or that observable value.

Evidently the theoretical model resulting from this novel approach is quite unlike that with which the classical or Newtonian theory has made us so familiar and which is so easy to visualize. But it is just this yearning for a theory that can be readily pictured in terms of diagrams and mechanical models which is apt to lead us astray; and Heisenberg rightly claims that the special merits of his theory spring from the fact that it is purely symbolic. His warning might well be passed on to those enthusiasts who seek to interpret the workings of the mind by constructing or designing electronic automata, which will, in their view, reproduce the essential features of human or animal behaviour: " inferences ", he says, " that are based on mechanical models are as misleading as they are seductive "[1].

To render Heisenberg's principle plausible to the ordinary reader, most popular expositions point out that, quite apart from the optical limitations of the microscope, any attempt to observe the exact position of an electron—e.g., by directing a beam of light upon it—would inevitably change its position, since a photon itself has a linear momentum (the ' Compton effect '): the difficulty is much the same as that encountered by the psychologist in attempting to introspect a feeling or to demonstrate ' free will '; his own efforts tend to modify the feeling and to disturb the process of willing. But in physics this type of explanation leads the reader to infer that, if only some more perfect method of observation were available, the uncertainty would be eliminated. All such inferences, however, are quite mistaken. The formal proof is purely algebraic, and makes no reference at all to the methods of measurement: (see appendix, p. 93). Indeed, once we accept the fundamental notion of an indivisible quantum, the indeterminacy becomes almost intuitively obvious[2]. Classical physics and the traditional notion of rigorous determination are based on observations and experiments concerned solely with what may be called man-sized phenomena, i.e., with objects whose magnitudes and velocities are within the range of our own sensory processes—notably vision: in such cases the uncertainties of quantum phenomena and the modifications required by relativity may be safely neglected. But in microphysical investigations, and in dealing with velocities such as obtain inside the atom or in the galactic systems, they acquire a paramount importance.

2. The majority of textbooks prefer to adopt a rather different mathematical model, due in the main to Schrödinger—possibly because, as Eddington once caustically remarked, " of the three alternative methods, it seems to be the only one which is simple enough to be

[1] The fault of all visual models is that they import into the abstract logical structure of the theory latent and often unrecognized postulates that are not only superfluous but actually in conflict with the facts. As Hamilton long ago realized, it was this that produced the apparent conflict between the two theories of light: the pictures are incompatible, though the ' algebra ' is not. Hence as he explains in his *Theory of Systems of Rays*, he deliberately determined to " replace diagrams by functions, and figures by formulae ". Niels Bohr himself has acknowledged the breakdown of his own mechanical model, which " pictured the atom as a small solar system ", and has readily accepted the newer theory as developed by Heisenberg and Schrödinger (cf. N. Bohr, *The Atomic Theory and the Description of Nature*, pp. 109f.).

[2] Mr. Clarke alludes to Einstein's argument (put forward during a debate on Heisenberg's principle at the Brussels Congress in 1930) that the apparent uncertainty resulted solely from the imperfection of existing human apparatus and techniques. To prove his point, Einstein described one of his ' thought experiments ' in which he sought to demonstrate how, in theory though not in practice, an instrument might be constructed to measure energy and time with perfect exactitude. Niels Bohr, however, during the following session, pointed out an omission in the reasoning, and showed that, even with ideal appliances, the uncertainty would still appear. Einstein accepted the criticism, though he still pinned his faith to a deterministic postulate.

misunderstood "[1]. In Schrödinger's model, p and q are not matrices but operators. The surprising result is that Schrödinger's model and Heisenberg's model, though so different in their starting-points, lead to virtually the same conclusions; and, as Schrödinger himself later showed, the one can be transformed into the other algebraically, much as the results of one mode of factor analysis can be transformed into those of another by ' rotating the factors '.

His general mode of approach is well worth studying, particularly as the future factorist may, in many of his problems, find it advantageous to follow the physicist, and substitute dynamic ' operators' for his static ' matrices ', thus changing some of his own fundamental postulates. For that reason it may be helpful to indicate the essential characteristics of Schrödinger's procedure by setting them out, roughly and briefly, in axiomatic form. The following are the most important of his initial postulates. (i) To every observable there corresponds a mathematical ' operator '; (ii) to every state of a physical system there corresponds a ' state function '; (iii) in any assembly of physical systems which include a dynamic variable open to observation the only possible values for that variable are the ' Eigen-values ' of the corresponding operator (i.e. the ' latent roots ' of the corresponding matrices); (iv) in such a system the possible values for the state functions may be obtained by solving an equation derived from the classical (Hamiltonian) energy equation by substituting differential operators for p (momentum) and E (energy).

From these postulates several instructive corollaries may readily be deduced. For example, (*a*) when the potential energy of the system is not explicitly time-dependent, the differential equation can be reduced to a simpler form known as the ' stationary state equation '[2]; (*b*) particles in stationary states act as though guided by a wave in multi-dimensional space whose amplitude and wave-length can be computed from the foregoing equation, and a probability distribution can be derived for the observable variables which will itself be independent of time; (*c*) when a physical system changes from one state to another, the change may be expressed by calculating certain ' transition probabilities '[3], which in turn depend on the ' matrix components ' of the wave functions for the two consecutive states. (The more important equations, and some indication of the methods of deducing the main results, are given in the appendix below, pp. 85–93.) Once again it follows that there can be no unique or determinate relation between one state and the next: the relations are essentially relations of probability.

3. The methods developed by Heisenberg and Schrödinger are appropriate only to cases that involve relatively low speeds. When the speeds are high, certain modifications are required to bring the theory into harmony with the principle of relativity. This was

[1] Schrödinger's own comment is worth quoting. " Owing partly to the difficulties involved in the kind of algebra employed and the absence of any possibility of visualization (*Anschaulichkeit*), I myself felt somewhat repelled (*absgescheckte, um nicht zu sagen abgestossen*) by his [Heisenberg's] theory: my own theory was prompted by de Broglie and certain brief but far-reaching remarks of Einstein's " (' Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der Meinen ', *Ann. d. Physik*, LXXIX, 1926, p. 735, footnote).

[2] The concept of a ' stationary state ' is perhaps best explained by quoting the first of Niels Bohr's ' postulates ': " An atomic system possesses a certain manifold of states, the ' stationary states ', to which corresponds in general a discrete sequence of energy values and which have a peculiar stability: every change in the energy of the atom must be due to a ' transition ' from one stationary state to another " (*loc. cit.*, p. 31). The mathematical theory of stationary states provides the basis, not only for quantum mechanics, but also for the greater part of quantum chemistry.

[3] Let me add that the concept of ' transition probabilities ' is by no means a monopoly of the quantum theorist, as Mr. Jackson seems to suppose: cf. M. S. Bartlett, *Stochastic Processes*, 1955, pp. 38 and 85f., and J. E. Moyal, ' Stochastic Processes and Statistical Physics ', *J. Roy. statist. Soc.* B. XI, 1949, pp. 150f. The student will find it instructive to compare Bartlett's matrix method of treating such probabilities with the matrix methods of Temple (*General Principles of Quantum Theory*, pp. 28–32).

effected by the more complicated theory proposed by Dirac[1], which postulates for a single electron, not one wave function, but four. Nevertheless, as is now becoming clear, none of the pre-war theories is adequate to account for all the known facts. The bombardment of the nucleus with particles of very high energy (100 million electron volts or more) has disclosed entirely unexpected results; what were supposed to be ' fundamental ' particles such as protons and neutrons, may actually be transformed one into the other; new particles (mesons, hyperons, anti-protons, anti-neutrons, etc.), mostly short lived, have made their appearance; and certain well-established principles—e.g. the ' conservation of parity '—seem to have broken down. The latest types of accelerators are beginning to suggest that the proton itself has a structure, and so seem to be revealing ' worlds within worlds ' even at the subatomic level. Current theories, however, were based on the assumption that each fundamental particle could be treated as a mathematical point, not as something occupying a finite region of space; and, when applied to problems which involve extremely small distances (10^{-13} cm or less), they yield impossible results: e.g. the values predicted for such properties as charge and mass tend towards infinity.

To meet these difficulties, Heisenberg, working in collaboration with Pauli, has endeavoured to reconstruct the earlier theories. Their starting point apparently is an equation of Dirac's, and the wave-function of Schrödinger is retained. But the ultimate aim is to produce a single comprehensive ' law ' or formula from which these novel phenomena may be satisfactorily deduced. These further developments, however, need hardly trouble us here. For our purpose it is sufficient to note that once again the principle of causality is explicitly discarded: " Nature ", says Heisenberg, " to a large extent behaves unpredictably "[2].

Let us now inquire what light, if any, is thrown on the problems of the psychologist by all these new discoveries and theories. Mr. Jackson asks whether my proposal to adapt the methods of the quantum theorists was " based merely on analogy " or whether I believe that " the neurological processes underlying human behaviour are of a quantum type ". The actual evidence does not, I think, at present warrant us in claiming more than that *some* neurological processes are of a quantal type; but with this change in wording, my answer would be ' both '. Some of the more obvious implications for psychology have already been stressed by quantum physicists themselves; and one object of the papers[3] that Mr. Clarke and Mr. Jackson now criticize was to bring such pronouncements to the notice of psychologists, who, in this country at any rate, appear so far to have overlooked their relevance. Bohr, for instance, writes: " The lessons we have learned from the discovery of the quantum of action open up to us new prospects which may perhaps be of decisive importance in discussions of the processes of living organisms in our picture of the world . . . The analogies between certain fundamental features of the quantum theory and the processes studied by psychology will not merely make it easier for us to

[1] P. A. M. Dirac, *Principles of Quantum Mechanics*, (1930). For the ordinary momentum of classical dynamics, $p = mv$, he substitutes a more general momentum $p + ek$, where e = the charge on the electron and k = the electromagnetic vector potential. His theory has two marked advantages: the equations retain their form under the Lorentz transformation (as required by the special theory of relativity), and they give a reason for the strange ' spin ' that characterizes the electron. A striking corollary was the existence of a positive counterpart to the negative electron, namely, the positron. Two years later this was experimentally verified by C. D. Anderson while studying cosmic radiation with the aid of the cloud chamber.

[2] My information is derived chiefly from the report of an address given by Heisenberg during the meeting of the division of theoretical studies of the European Organization for Nuclear Research at Geneva on 4 March of the current year.

[3] Most of the following examples were included in the papers cited; for reasons of space they were omitted from the printed abridgements except for references in footnotes. After all, my chief purpose was not so much to demonstrate any concrete conclusions of my own, but rather to make out a prima facie case for closer attention and research on the problems indicated.

adjust ourselves to the new situation in physics, but (it is perhaps not too ambitious to hope) may prove of value in our endeavours to obtain a comprehensive understanding of the more subtle problems of psychology ". Much the same points have been stressed, in even greater detail, by Schrödinger[1].

Since the quantity designated h is extremely small, the discontinuities and uncertainties which it measures may, as we have seen, be safely disregarded so long as we are dealing with macrophysical phenomena. But when we turn to physiological processes which depend on structures too minute to be visible even in the microscope, such a policy is likely to lead us astray. Darlington, for instance, estimates the volume of a gene (a highly complex molecule) as about that of a cube whose edge is about 300×10^{-10} m—a length that must be barely 100 atomic distances. The processes at the neural synapses apparently take place within a region whose volume is smaller still.

One special field in which the methods and results of quantum theory are particularly relevant is that of sensory discrimination. Discrimination depends on neural processes of an all-or-none type; and this itself suggests that, not only the absolute threshold for a sensory stimulus, but also the just perceptible differences, will at a given instant be finite and specific. The inference seems to be fully borne out by von Békésy's experiments on differential sensitivity to the intensity of sounds (G. v. Békésy, ' Über das Fechnersche und die Theorie des Hörens ', *Ann. Physik*, VII, 1930, pp. 329–359; cf. *id.*, ' The Recruitment Phenomenon and the Difference Limen in Hearing and Vibration Sense ', *Laryngoscope*, LVII, 1947, pp. 765–777). Stevens and others have confirmed his results, and obtained values which seem to reveal more precisely the size of the ' neural quantum .' Stevens indeed goes on to argue that in any case a quantal hypothesis provides a far better model for the psychologist than the phi function of gamma; and he regards the latter as merely a special case of the quantal hypothesis: (cf. S. S. Stevens *et al.*, " Theory of the Neural Quantum in the Discrimination of Loudness and Pitch "; also G. A. Miles and W. R. Garner, " Implications for a Quantal Theory of Discrimination ", *Amer. J. Psychol.*, LVII, 1944, pp. 451–467).

Similar problems arise in dealing with visual thresholds. Hecht, for example, has found that for minimum visibility the energy actually reaching the retina may be as small as 5 to 14 quanta spread over an area containing about 500 rods; this means that on an average only a single quantum of light needs to be absorbed by each of 5 to 14 rods (S. Hecht *et al.*, " Energy at the Threshold of Vision ", *Science*, XCIII, 1941, pp. 585–7).

There are many other peculiarities in the phenomena discussed by the quantum theorist which display an analogy—it may be merely a formal analogy—with phenomena encountered in psychology. In the field of volition, for example, as James and McDougall have repeatedly emphasized, " the paradox of willing lies in the fact that a decision based solely on ideas and ideals, and therefore weak, is able to overcome barriers, obstacles, and cravings far stronger than itself, acting often in the line of greatest resistance, in a way that contradicts all the analogies from dynamics ". This is surprisingly similar to what has been called " the chief paradox of quantum physics ": in classical mechanics, according to the law of the conservation of energy, a particle, whose energy is less than the potential energy of a barrier in its path, could never overcome that barrier; on the other hand, the study of the disintegration of radium by alpha decay and of the transference of an electron from one atom to another reveals that, " contrary to all expectation, the particle, after an interval of time, the length of which is a matter of chance, is able to break through the barrier as though through a tunnel " (Heisenberg, *l.c.*, p. 121). The theory of ' wave mechanics ' yields an intelligible account of this singular phenomenon; and an equation, formally analogous to Schrödinger's, seems to furnish a much better mode of describing volitional processes than James' oft-quoted equation (*Principles*, II, p. 549).

[1] For a non-technical account see his lectures entitled *What is Life ? The Physical Aspect of the Living Cell* (1955), esp. pp. 46f.

But above all, the results of quantum physics suggest an entirely new conception of matter. We have henceforth to think of it, no longer as a passive vehicle, whose first law is a law of inertia, but as something essentially and perpetually active. Consequently the methods to which such conceptions have led should be of special interest to those psychologists who are no longer content to regard the brain as a passive transmitter of impulses, waiting for some external ' stimulus ' to provoke it into a predetermined ' response ', but as an organ which itself is continually active, and indeed an organ whose spontaneous activity is vibratory.

Eddington has more than once pointed out that the quasi-statistical techniques of the quantum physicists are by no means restricted to problems arising within their somewhat specialized spheres; as he remarks, " one of the chief difficulties over accepting it as a theory of the atom " has been " the almost universal applicability of the procedure ". To begin with, we may note that the concept of a ' state function ', and the ingenious methods evolved to deal with it, could be extended to almost any dynamic system. It would be quite practicable to construct such a function for the solar system, and to use it for describing the motions of the planets, for making predictions about their future, and for computing the amount of uncertainty involved. No doubt, an ' assembly ' of such systems would be far from easy to procure, and little would be gained to compensate for the increased difficulty of the calculations. On the other hand, a model of this type could readily be adapted for describing human behaviour; and ' assemblies ' of individuals and of groups could unquestionably be subjected to analysis along these lines with a considerable addition to knowledge.

The formal analogies are often striking. The phraseology used by Heisenberg, for instance, to describe the ' stationary ' but active states of an atom, and to explain the way such a system is thrown into an ' excited state ' by a disturbing stimulus from outside, then reacts, and so either reverts to its original ' fundamental state ' or passes into a different ' stationary state '—all this (if we did not know the context) might easily be an account of the way an individual organism or a human being behaves before, during, and after it is ' disturbed ' by its environment. To a psychologist, indeed, the details of the account are curiously reminiscent of the theory of ' vital series ', put forward by Avenarius and developed by Stout[1]. In recent factorial studies, the mental ' states ' of individuals, and the transition probabilities for a possible passage from ' one ' state into another, have already been investigated by the aid of a matrix algebra not unlike that of Heisenberg, and Eigenvalues and Eigen-functions have been computed to express the results in canonical form: (see *Psychometrika*, III, pp. 151f.). For more complex problems the technique developed by Schrödinger, which uses operators instead of matrices, seems to offer a particularly convenient method of attack.

I think then we may justly claim that both the methods and the results of quantum physics deserve far more attention from psychologists than they have hitherto received. What further additions to knowledge such efforts will yield only actual trial can show. Meanwhile, one important conclusion stands out beyond dispute. The brain or central nervous system can no longer be treated as a ' determinate machine ' analogous to a man-made mechanism: its mode of operation must be essentially probabilistic.

Mr. Clarke is quite right in saying that the doctrine he upholds—" the doctrine of strict scientific determinism "—would be " in sharp conflict with our seeming sense of volitional responsibility "; and he himself contends that, since psychology cannot abandon the former without ceasing to be scientific, we must give up what he supposes to be " the mere dogma of a free or capricious will ". To this my answer would be that there are not two but three courses open to us: (i) we may maintain, as Mr. Clarke does, that our "sense of volitional responsibility " is both " an illusion and a dogma "; (ii) we may maintain, as James and other indeterminists have done, that what Mr. Clarke calls " the imperative

[1] G. F. Stout, *Analytic Psychology*, I, pp. 151f., 268f., and for a succinct statement of Avenarius' theory, see. J. Petzoldt, *Vierteljahresschrift f. wiss. Philosophie*, XVIII, ii, pp. 203f.

nature of the principle of causality " is a dogma and possibly an illusion; or (iii) finally, we may try to show that, when the two sets of phenomena are correctly interpreted, without invoking ' dogmas ' of either type, there is in point of fact no conflict between them. Far from being a ' mere dogma ', our ' sense of volitional responsibility ' (for such it must surely appear to an unbiased critic) is an intuitive datum, which itself demands scientific examination. It implies not capricious volition, but rational volition. Equally I should hold that the so-called principle of causality implies a belief, not in a deterministic compulsion, but in a rational sequence underlying phenomena: in other words, that there are intelligible ' grounds ' for the ' consequents ' we have undertaken to study. In any case, the concept of causality is not a principle whose nature is ' imperative '; it is simply a postulate, which it may be convenient to adopt for some scientific models or for some types of exposition, but inconvenient and certainly unnecessary for others.

Mr. Clarke assures us that " science has been based on determinism ". But the most he is entitled to claim is that scientists (or many of them) have based their investigations on the assumption of determinism as a working hypothesis[1]; and to them, at least in practice, ' determinism ' has merely meant, not temporal consequences enforced by antecedent causes, but logical consequences deducible from some simple relation, preferably a soluble mathematical equation. In the past, when they were dealing with macrophysical phenomena, the equations were ' determinate ', and the inferences that were based on them could be treated as exact and certain. In the new field of microphysics this is no longer permissible; and so the great Machine Theory of the Universe collapses.

[1] Mr. Clarke quotes Hebb's remark about ' complete causation '; but, if he reads on, he will find that Hebb later observes: " actually of course this is a working assumption only " (*loc. cit.*, p. xiii). No doubt, as Mr. Clarke says, " any proposal to abandon determinism would be revolutionary "; but then so are the new discoveries that have given rise to such proposals.

Nevertheless, it is only fair to recognize that one or two physicists have recently attempted something like a counter-revolution. David Bohm, for instance, has suggested (*Causality and Chance in Modern Physics*, 1957) that the " apparently random fluctuations at the quantum level " may be determined by " hidden variables at the subquantum level ", much as the erratic ' Brownian movements ' of particles of pollen at the microscopic level are determined by the constant bombardment of molecules at an infra-microscopic level. However, we know a good deal about the activities of molecules, but nothing at all about ' hidden variables ' at the ' subquantum level '. It is, of course, conceivable that the new accelerators—the synchrotons and cosmotrons now being built—may disclose such evidence. But on the whole the tendency of the most recent discoveries has been to increase the importance of a probabilistic treatment rather than to diminish it; and, with one or two notable exceptions, Bohm's proposals have so far found little favour in the eyes of most quantum physicists. Professor Rosenfeld, for example, in a review of the book, remarks that " it is impossible to discern the slightest foundation for the surmise that a return to a deterministic mode of description could throw any light on the phenomena " (*Nature*, No. 4610, 1958, p. 658).

Bohm's main quarrel, however, seems to be not so much with the philosophy of indeterminism as such but rather with what he calls ' mechanism ', whether determinist or indeterminist, i.e., with the attempt to reduce the qualitative differences of the universe to differences exclusively of quantity. And, since he insists on the ' qualitative infinity ' of the universe and on the possibility of an unending series of levels, his postulates afford little encouragement to the thoroughgoing determinist. Indeed, at one point he himself talks of " chance fluctuations as an inseparable part of the mode of being of things ".

Mr. Clarke's comparison of the probabilities that occur in quantum theory to the probabilities that arise in drawing cards is open to the same objections. With the cards we know that the unpredictable order is due to the complex rearrangement introduced by the dealer who " shuffled the pack ": but we know of no dealer who is responsible for the unpredictable behaviour of electrons. All such speculations are plainly based more on metaphysical preferences than on the available data. Furthermore, they seem to be negatived at the outset by a well known theorem of von Neumann, in the effect that the behaviour of events at the atomic level is such that no precisely specifiable distribution of ' hidden variables ' could possibly account for it. Hence, as Kauzmann and others have observed, " most scientists today accept the uncertainty principle as the expression of an inherent characteristic in the nature of things " (W. Kauzmann, *Quantum Chemistry*, 1957, p. 240).

Even if we admit that Bohr, Eddington, Heisenberg, Schrödinger, and others who adopted similar views may have pressed their conclusions too far, it still remains true that the principle of determinism cannot be a *necessary* element in scientific explanation: else we should not find so many eminent scientists ready to dispense with it. The substitution of probability for rigorous determinism therefore springs, not " from a preference for a vague irrationality over causality ", but rather from the view that the phenomena of the universe may be the resultants *partly* of principles implying a rational order and *partly* of random or unorderly elements of contingency. Surely, in a world where much that is logically possible has not come into existence and where all that does exist is not logically necessary, a theory of *limited determinism* is not unreasonable.

My answer then to Mr. Clarke's main question is briefly this: the so called ' principle of indeterminacy ' does not mean, and does not entail, ' indeterminism ' in the psychologist's sense; but it does remove one of the main objections to it, should we have reason to adopt it on other grounds. Let me add that I heartily agree with him that " psychology should not content itself with statistical studies alone ": " experimental studies " are equally essential; and the two should go hand in hand. Certainly one of the papers that he criticizes (R. J. Audley and A. R. Jonckheere, ' Stochastic Processes and Learning Behaviour ', this *Journal*, IX, pp. 87–94) was purely statistical; but its object was merely to introduce and defend the methods the investigators are already using to analyse results obtained in actual experiments on learning in animals. The paper by Miss Foley and myself included experimental data (see pp. 54 and 60f.): these were fitted by applying matrices for ' transitional probabilities ', deduced from certain postulates. I may add that in more recent experiments, in which larger numbers have been tested, we have found that, by substituting ' operators ' for matrices, and calculating ' state functions ' for the ' assemblies ', we secure, if anything, a still closer fit, and (once the algebraic procedure has become familiar) an even simpler procedure for the necessary calculations. We hope to report the final results in a future article, to which part of the foregoing discussion may be regarded as a preamble.

CYRIL BURT

APPENDIX: WAVE MECHANICS AND PROBABILITY FUNCTIONS[1]

1. *Oscillations in General.* The phenomena of wave-motion are met with at every turn in ordinary life—in the ripples on a lake, the swing of a pendulum, the bobbing of a

[1] I have added this somewhat technical appendix in response to the request made by Mr. Jackson and other correspondents for " some clearer indication of what the mathematical arguments really are ". My aim has been to pick out the topics that seem most relevant to the problems of the psychologist, and to start from a type of wave motion with which even the Arts student becomes acquainted during his elementary courses in the psychological laboratory. In so doing I have no doubt committed blunders and overlooked points of importance. Others more expert will, I hope, correct these defects. But my chief object is to persuade the research-student to dip into the literature of the subject himself. For the psychologist the most instructive publications are the original papers of Heisenberg and Schrödinger: there he can see the methods actually taking shape; and, since I have also been asked to suggest one or two introductory textbooks, may I recommend (again rather with the special interests of the psychologist in mind), first, S. Dushman, *The Elements of Quantum Mechanics* (New York, Wiley, 1938), and secondly W. Kauzmann, *Quantum Chemistry* (New York, Academic Press, 1957) ?

In our further correspondence on the problems here discussed, Mr. Jackson has pointed out that many students of psychology are nowadays unacquainted with the detailed psychology of auditory sensation: " some recent textbooks (e.g., P. McKellar's *Textbook of Human Psychology*) do not mention sound at all ". At his suggestion therefore I have amplified the somewhat elementary explanations in the introductory paragraphs. Readers to whom this is already familiar may take as their starting point the classical equation of the electromagnetic theory of light (cf. eqn. 25).

My colleagues, Dr. E. A. Power and Mr. L. A. Wigglesworth (of the Departments of Mathematics and Physics, University College, London), have been good enough to read the whole appendix in draft; and I am deeply grateful to them for several emendations and suggestions.

weight suspended by a spring, and " the sound of the cornet, flute, harp, dulcimer, and all kinds of musick ". In studying visual sensations the psychological student learns of electro-magnetic waves, and, in studying the brain, of the rhythmical waves of electric potential recorded by the electroencephalogram. To analyse such phenomena with scientific precision the stock mathematical procedure is to set up a differential equation which will describe, at least in all essentials, the motion observed, and then to solve this equation by means of integration, using the initial or boundary conditions to determine the constants that inevitably appear. In every case the mathematical model proves to be formally the same. And to the student of psychology the general formula is probably most familiar from his studies of the physical stimuli for musical tones.

The simplest type of vibratory motion is that obtained with a tuning fork or electronic oscillator. This yields what the musician calls a ' pure ' (or nearly pure) tone. When a prong of the fork is displaced, a tension is set up which tends to bring it back to the position of equilibrium and actually swings it over in the opposite direction, so that a regular oscillation results. In an ideal model, the tension of the restoring force, and therefore the magnitude of the acceleration, will be proportional to the amount of displacement of the vibrating portion from the centre of oscillation, and will be directed towards that centre. Thus, if y denotes the displacement at any time t, then the acceleration will be

$$\frac{d^2y}{dt^2} = -\omega^2 y, \qquad \ldots (1)$$

where ω^2 is the constant of proportionality[1]. Since this is a differential equation of the second order, the general solution will require two constants. The student should recognize that, owing to the relations between trigonometrical and exponential functions, the solutions may be written in various forms, e.g.

$$y = a_1 \exp(i\omega t) + a_2 \exp(-i\omega t), \quad \text{where } i = \sqrt{-1} \qquad \ldots (2)$$

$$y = A_1 \sin \omega t + A_2 \cos \omega t, \qquad \ldots (3)$$

$$y = A \sin(\omega t + \phi_1), \qquad \ldots (4)$$

or, since $\cos\theta = \sin(\theta + 90°)$, $y = A\cos(\omega t + \phi_2)$. In all these equations the constants of integration, a_1 and a_2, A_1 and A_2, or A and ϕ_1, have to be determined by two initial conditions, independently given. Since $\sin\theta$ and $\cos\theta$ are periodic functions with a period of 2π, the motion is ' periodic '; and, since it can be represented by a sine curve, it is called ' sinusoidal ' (or sometimes ' pendular ' or ' simple harmonic ').

Evidently the to-and-fro motion might be regarded as the projection, on a given straight line, of a radius-vector whose length is A and which rotates with a constant angular velocity of ω rad/sec: if the motion starts with the vector lying along the given line, the angle ϕ will be zero. Thus ϕ represents the 'phase angle', and A the 'amplitude' of the wave. If T sec. is the time or ' period ' for one rotation, then $\omega = 2\pi/T$; and the frequency of the

[1] In view of subsequent developments, let us note that the solutions of linear equations of this type are equivalent to determining what are known as the Eigen-values (' proper ' or ' characteristic ' values) and the Eigen-functions of the operators involved. These terms may be defined as follows. Let R be any operator and y any function: then, if

$$Ry = \rho y,$$

y is said to be an ' Eigen-function ' of the operator R, and ρ an ' Eigen-value '. The psychological factorist will be familiar with this type of equation from problems in which R is a correlation matrix, the values for ρ its ' latent roots ' or ' factor variances ', and y the corresponding ' latent vectors ' or ' factor loadings ' (duly normalized): the latter may be regarded as direction-cosines for a vector, which, when operated on by R, is changed only in magnitude, not in direction; and, if R is symmetric, these vectors are mutually orthogonal. Or again to take a somewhat simpler example, if R denotes the operator d/dx, then $\exp(\omega x)$ will be an Eigen-function of R, with an Eigen-value of ω, because $(d/dx)\exp(\omega x) = \omega\exp(\omega x)$: or once again, if R denotes the operator d^2/dx^2, then $\sin\omega x$ will be its Eigen-function with an Eigen-value of $-\omega^2$. This may serve to explain the use of these two technical terms in the paper criticized by Mr. Jackson (Burt, *Psychometrika*, III, pp. 151f.).

oscillation or ' wave ' will be $\nu = 1/T = \omega/2\pi$. We thus obtain the simplest possible expression for a harmonic wave,

$$y = A \sin 2\pi \frac{t}{T} = A \sin 2\pi\nu t. \quad \ldots (5)$$

This is the formula commonly given in textbooks of psychology. Impinging on the ear, an atmospheric ' wave ' of this shape will produce a ' pure tone ': its pitch will depend on the frequency, ν, and its loudness on the square of the amplitude, A.

2. *Travelling Waves.* If a cord is fixed at one end, and the other end be given a vertical jerk, a wave-like displacement will pass along the cord; the tighter the cord, the more rapidly will the wave travel from the source of the disturbance. Let x denote distance, measured along the cord when at rest, m the mass of the cord per unit length, and F the tension. At successive moments of time, the amount of displacement, y, will be different at different points along the cord. From the ordinary Newtonian laws of motion, we can deduce

$$F\frac{\partial^2 y}{\partial x^2} = m\frac{\partial^2 y}{\partial t^2}, \quad \ldots (6)$$

or

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}, \quad \ldots (7)$$

where $v \equiv (F/m)^{\frac{1}{2}}$ denotes the velocity of the individual waves. This latter formula—the ' wave equation '—gives a complete description of wave motion with one degree of freedom, i.e. a motion travelling in one dimension (x) only. The equation remains valid provided F does not depend on x (though m may): both m and F may depend on t. When F and m are both constant, the solution of eqn. (7), expressed in the trigonometrical form, is commonly given as

$$y = A \sin 2\pi(x/\lambda - \nu t) \quad \ldots (8)$$

where ν as before denotes the frequency and $\lambda = v/\nu$ denotes wavelength.

When, however, v^2 is a constant, the two variables, x and y (distance and time), are separable. Accordingly, let us write $y = f(x)g(t)$, where $f(x)$ is a function of x only and $g(t)$ of t only. Substituting in eqn. (7), we obtain the ' ordinary ' differential equations

$$\frac{1}{f(x)}\frac{d^2 f(x)}{dx^2} = \frac{1}{v^2}\frac{1}{g(t)}\frac{d^2 g(t)}{dt^2} = -k^2 \quad \text{(say)}.$$

Putting the solutions for the two equations in exponential form, we have

$$f(x) = A \exp(ikx) + B \exp(-ikx) \quad \ldots (9)$$

$$g(t) = C \exp(ikvt) + D \exp(ikvt), \quad \ldots (10)$$

where A, B, C, and D are arbitrary constants to be determined by the initial or boundary conditions.

3. *Stationary Waves.* Now let us turn to the case in which both ends of a uniform cord are fixed, as with a piano wire or the string of a harp or violin. F and m will then be constant. Let L denote the length of the string or wire. Then, when $x = 0$ or L, $f(x) = 0$; hence $A + B = 0$; and when $x = L$, $A[\exp(ikL) - \exp(-ikL)] = 0$, i.e. $\sin kL = 0$. Thus a physically significant solution is obtainable only for certain discrete values of k, namely, when $k = n\pi/L$, and n is *any whole number* (i.e. 1, 2, 3, . . .). Much the same arguments hold good for $g(t)$, and so for the constants in eqn. (10). It will be seen that the values obtained for k are the Eigen-values of the operators involved (see footnote 1, p. 86 above), and the corresponding values for the functions $f(x)$ and $g(t)$ are the Eigen-functions.

And now, changing back to trigonometrical form, we may evidently write, as a final solution of eqn. (7),

$$y = f(x)g(t) = A \sin n\pi x/L \,.\, \sin 2\pi n\nu_0 t, \quad \ldots (11)$$

where ν_0 denotes the fundamental frequency.

If we assign a constant value to x, the expression containing t will describe the motion of a particle at the point so specified. If, on the other hand, we assign a constant value to t, then the expression

$$y = A \sin n\pi x/L \qquad \ldots (12)$$

will describe the form of the wave at the time specified. Equation (12) is ordinarily termed the ' normal mode equation '. It contains only the coordinate (x), whereas the ' wave equation ' (eqn. (7) and its solution) contained the time (t) as well. Accordingly, waves described by equations such as eqn. (12) (e.g. waves that do not travel beyond the limits of the cord) are commonly referred to as ' standing ' or ' stationary ' waves.

The appearance of the integral multiplier, n, corresponds with the fact—familiar to most psychological students of psychology from their laboratory work—that, when a stretched string is plucked exactly in the centre, it usually vibrates as a single whole ($n = 1$, giving the fundamental note). If we interfere with this fundamental mode of vibration, e.g. by touching the string lightly in the middle with a brush, it then vibrates in halves giving the octave ($n = 2$). Plucked at one-nth of the total length (where n is any whole number), it will vibrate in smaller sections, giving other overtones or ' harmonics '. When bowed, it commonly vibrates with a still more complex wave, formed, as it were, by the superposition of the constituent waves[1]. But the main point to notice is that the scale of vibrations that can be elicited from a string of fixed length and constant tension forms, not a continuous sliding scale, but a discrete series: the string will vibrate, not with any or every conceivable frequency, but only with those particular frequencies that form whole multiples of the fundamental frequency. When, for instance, a violinist has tuned his G-string to give a note of 192 vibrations per second, he cannot get a note of 200 vibrations except by stopping off the string with his finger, and so reducing its length: but he can easily produce a note of $192 \times 3 = 576$ vibrations (d'). Frequency (ν) and wavelength (λ) are inversely related. If the frequency is $\nu_n = n\nu_0$, then the wavelength will be

$$\lambda = \lambda_0/n = 2L/n,$$

where ν_0 and λ_0 are the frequency and the wavelength of the fundamental note. Hence there will be $(n-1)$ ' nodes ' (or points of zero displacement) between the two fixed ends of the string.

In much the same way it may be shown that, whenever a wave-like motion is restricted to oscillations within a limited space, then the only possible frequencies that it can exhibit will be a discrete and not a continuous series.

4. *Wave Mechanics.* According to the ordinary laws of classical physics, we should expect the electrons revolving round the nucleus of an atom to emit a continuous radiation. Their emission spectra would therefore be continuous; and their activity would be progressively diminished or damped; hence they would spiral round the nucleus, and eventually fall into it. Experimental observations show that none of these results occurs. To meet this discrepancy between fact and theory, Bohr suggested a couple of *ad hoc* postulates: (1) that an electron can, at least for an appreciable time, circle round the nucleus at certain

[1] This last statement follows because, with a linear differential equation, the sum of any set of particular solutions is itself a solution. The superposition of such waves is called a ' Fourier synthesis '. The reduction of a complex wave-pattern to a series of simple harmonic waves is called a ' Fourier analysis' ; and the expressions so obtained serve to describe an astonishingly wide range of complicated changes. The resulting components are orthogonal. Hence this mode of analysis exhibits many striking analogies with more familiar types of ' factor analysis '.

It should be added that, throughout the foregoing discussion, I have for simplicity ignored any possible loss of energy in the vibratory system. To allow for such ' damping ', a further term must be introduced into the initial differential equation. In psychological processes it often happens that the opposite type of progressive change appears: the wave-like function acquires increasing definiteness, as though the system were accumulating energy instead of losing it as in forced oscillations and resonance. This is observable in processes of growth and learning.

fixed distances without radiating energy, and that there is a series of ' permissible orbits ' in which this is possible; (2) that, when it emits radiation, it jumps abruptly to an orbit nearer to the nucleus, and when it absorbs radiation, it jumps outward, the change in energy, viz., $e^2/2r_2 - e^2/2r_1$, being in either case equal to $h\nu$, where r_1 and r_2 are the radii of the orbits, e the charge on the nucleus, and ν the frequency of the radiation. Thus the atom can exist only in one of a sequence of ' stationary states ', each characterized by a definite energy level.

De Broglie, as we have seen (p. 78), put forward, on quite general grounds, a more comprehensive hypothesis: he assumed that a material particle, such as an electron or photon, has its motion regulated as it were by an associated wave whose wavelength is $\lambda = h/mv$, where m denotes the mass of the particle and v its velocity. When the waves are confined within an atom, their possible frequencies, like those of a vibrating string fixed at both ends, will form a discrete series; and their wavelengths will then be

$$\lambda = \lambda_0/n,$$

where λ_0, as before, denotes the fundamental wave length and $n = 1, 2, 3, \ldots$ etc. Thus the properties of the ' stationary states ' of an atom can be interpreted as properties of ' stationary ' waves, i.e. of the ' normal modes ' of a system in harmonious vibration. With these assumptions the angular momentum of the revolving electron will be

$$mvr = nh/2\pi;$$

and its kinetic energy $T = \frac{1}{2}mv^2 = n^2(h^2/2m\lambda^2)$. Hence, like the frequencies, the energy levels of the electron will also form a discrete series; and the energies will be proportional to a sequence of whole numbers, while the radii of the corresponding orbits will be proportional to their squares.

These are precisely the results that Bohr had arrived at on the basis of the experimental data. Thus, as Heisenberg puts it, " an atom can ' vibrate ' in its normal or fundamental state, in which case it has no nodes, i.e. no levels at which the density of matter disappears " (this corresponds to the orbit of lowest energy): when ' stimulated ' by the impact of an incoming radiation, " it may vibrate in an overtone, i.e. in an ' excited ' vibration, in which there may be several nodes "; and it may ' react ' by emitting a quantum of light energy, and so " pass from a highly excited state to another of smaller energy, or back to its fundamental state "[1].

It was the purpose of Schrödinger's ' wave mechanics ' to investigate the formal properties of these stationary states in greater detail[2]. His basic equation is a partial differential equation, similar in type to the differential equations for waves of sound or light, and describing the manner in which the wave-pattern, produced by the motion of an intra-atomic particle in a given field of force, varies with the space coordinates. If the vibrating system is a single particle with three degrees of freedom, its ' state ' can be represented by the function $\Psi(x, y, z, t)$, where x, y, z denote Cartesian coordinates. This is known as the ' state function '; it is explicitly ' time-dependent ', and is commonly denoted by a capital psi. Ψ is thus analogous to y in the preceding analysis, where $y = y(x,t)$ represented

[1] W. Heisenberg, *Nuclear Physics*, 1953, pp. 36f. Bohr's theory of circular orbits accounts very neatly for the main lines of the hydrogen spectra, though not for the ' fine-line ' structure observable with high resolving power: these require the assumption (which in fact follows from the ordinary inverse square law) that the orbits should be elliptical and not circular.

[2] The essential features of Schrödinger's theory were described in a series of four papers on ' Quantisierung als Eigenwertproblem ' (*Ann. d. Physik*, LXXIX, 1926, 361–376, 489–527, LXXX, 85–138, 139–169). The demand for these papers was so great that a few months later they were reprinted, together with two other contributions and a useful rearranged summary of the contents, under the title *Abhandlungen zur Wellenmechanik* (Leipzig: Barth, 1927). I cannot resist quoting the comment of *eine junge Freudin* to Schrödinger (p. iii): " Gelt, das hätten Sie sich auch nicht gedacht, als Sie anfingen, dass dabei so was Gescheites herauskommen würde! ".

the ' state ' of a string or cord of indefinite length. The corresponding ' wave function ', describing a ' stationary state ', will be $\psi(x, y, z)$: it is no longer ' time-dependent ', and is commonly denoted by a lower case psi. If the system contains n particles, then the ' state function ' will be $\Psi(x_1, y_1, z_1, x_2, \ldots, z_n, t)$ with a similar modification for the wave function.

To simplify the problem we may begin by considering a single particle, whose mass is m moving in a field of force for which the potential energy U is a function of a single coordinate only, namely, x. In its simplest form, then, our initial assumption may be written

$$\frac{\partial^2\Psi}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\Psi}{\partial t^2} \qquad \ldots (13)$$

where as before $v = \nu\lambda$. This is a ' time-dependent equation ', and is analogous to eqn. (7) above. In trigonometrical terms the simplest solution would then have the same form as in eqn. (8), namely,

$$\Psi = A \sin 2\pi(x/\lambda - \nu t) \qquad \ldots (14)$$

Now, in accordance with Bohr's quantum conditions, we may take the frequency to be given by $\nu = E/h$; and in accordance with de Broglie's postulate we shall assume that the wavelength is given by

$$\lambda = h/p = h/mv = h/\sqrt{\{2m(E-U)\}}, \qquad \ldots (15)$$

where p denotes linear momentum, m the mass of a single particle (e.g. an electron in a hydrogen atom), v its linear velocity as it revolves in its circular orbit, E the total energy, and U the potential energy at any point expressed as a function of the coordinates. Hence we can rewrite eqn. (14) in the form

$$\Psi = A \sin\frac{2\pi}{h}(px - Et). \qquad \ldots (16)$$

In order to illustrate quite briefly the further developments of the theory, it will be better to change to the equivalent exponential form. Under certain conditions, e.g. when the potential energy does not depend explicitly on the time, it is possible (as with the ' stationary waves ' of a vibrating string, cf. eqn. (11)) to apply the method of separating variables, and put

$$\Psi(x, t) = \psi(x)\phi(t). \qquad \ldots (17)$$

Then, making the appropriate substitutions from eqn. (15), we obtain for the time function (as in eqn. (10))

$$\phi(t) = A \exp(-2\pi i/h)Et. \qquad \ldots (18)$$

However, since we are here primarily concerned with the ' stationary ' wave patterns, let us keep mainly to the coordinate-function. Making the appropriate substitutions once again, we find

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m(E-U)\psi}{h^2} = 0. \qquad \ldots (19)$$

This is Schrödinger's equation for the ' steady state '. One important feature should be noted. In the examples mentioned in sect. 1 the values of y (which in the case of sound, represents the longitudinal displacement of an atmospheric particle from its position of equilibrium) must necessarily be ' real '; in quantum theory ' complex ' values of ψ are allowable, and indeed almost essential. Since the operators are Hermitian, all the eigenvalues and therefore all the measurable quantities will be ' real ' numbers. Various cases arise according to the relative values of E and U: we need glance only at one or two special results.

When there is no field of force, $U = 0$; and eqn. (19) reduces to

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 mE\psi}{h^2} = 0. \qquad \ldots (20)$$

This, as Schrödinger says (see eqn. (29), p. 47, of his second paper), " is altogether the simplest conceivable instance of the wave theory "; and the solution may be written

$$\psi(x) = A\exp(ikx) + B\exp(-ikx), \qquad \ldots (21)$$

or, as he puts it (in trigonometrical form),

$$\psi(x) = \begin{matrix}\sin\\ \cos\end{matrix}\left[\sqrt{\left\{\frac{8\pi^2 mE}{h^2}\right\}}x\right], \qquad \ldots (22)$$

where k^2 in eqn. (21) denotes the expression under the radicle in eqn. (22) (*loc. cit.*, eqn. (30)). The values for $E(E_n$ say) can have physical significance only if $E_n = n^2h^2/8\pi^2 m$, where n is again *a whole number*, a result which is " in perfect accord with the results of quantization " (Schrödinger, *loc. cit.*, p. 48).[1]

When a field of force is present and the potential energy $U(x)$ varies, the deductions become more elaborate. In the case of a ' simple harmonic oscillator ' (cf. sect. 1 above) the potential energy is $\frac{1}{2}kx$; and the wave-function ψ then takes the values

$$\psi_n = C_0 H_n(\xi)\exp(-\xi^2/2) \qquad \ldots (23)$$

where $\xi = 2\pi x\{\sqrt{(\pi m v_0/h)}\}$, $C_0 = \{2/\sqrt{(\pi m v_0 h)}\}^{1/2}$, and $H_n(\xi)$ denotes the Hermite polynomial[2] of order n; and for the various energy levels, E_n, we obtain the discrete values $h(n+\frac{1}{2})$, where as before $n = 0, 1, 2\ldots$. It will be observed that, when $n = 0$, the expression ψ_n is the equation for the ordinary normal or Gaussian probability curve (Schrödinger, *Naturwiss.*, XXVIII, 1926, pp. 664f.).

The values obtained from the various solutions agree far better with spectroscopic observations and other data than any earlier theory: and they supply a key to the chemist's periodic table. Furthermore, they explain one of the most paradoxical phenomena encountered in sub-atomic investigations, namely, the way in which electrons are able to tunnel through barriers of potential energy (e.g. those due to the Coulomb forces arising from a large nuclear charge) when this is much greater than the energy of the particles; they even yield a so called ' transmission coefficient ' to express the probability that an electron will get through.

Waves of Probability. In the early popular discussions of quantum theory (with which alone psychologists seem familiar) rather a mystery was generally made of the ubiquitous symbol psi. Eddington suggested that we should regard it " as a fundamental indefinable of the theory "; others suggested that it described some amphibious entity that might be called a ' wavicle '. If, however, we revert to the analogy with the waves of sound or light, we may recall that the square of the amplitude, A^2, measures the intensity of the sound or light. Now, taking the value for psi given by eqn. (17), and rewriting the functions $\psi(x)$ and $\phi(t)$ in the simplest exponential form, we have, on multiplying Ψ by its complex conjugate,

$$\begin{aligned}\Psi\Psi^* &= AA^*\exp(ikx)\exp(-2\pi ivt)\exp(-ikx)\exp(2\pi ivt)\\ &= AA^* = |A|^2,\end{aligned} \qquad \ldots (24)$$

a ' real ' magnitude; and this may accordingly be regarded as expressing the intensity of the beam that crosses unit area per second, or, what amounts to the same thing, the number

[1] The acceptable values must be determined by the *boundary conditions*. For instance, an electron whose behaviour is restricted by the electrostatic field around a positively charged nucleus will behave like a particle in a ' box ': within the box-like region the potential energy (u) may be taken as zero; outside it, as infinite. When we are considering a single dimension only, the situation is analogous to that of a bead sliding along a wire with stops or barriers at each end, i.e., where $x = 0$ and $x = L$ (length of wire). Outside the box therefore $\psi = 0$. Hence the electron cannot appear outside the region enclosed by the barriers formed by the potential. In this case the problem is virtually identical with that of finding the ' normal modes ' of a vibrating string fixed at both ends.

[2] The Hermite polynomials will be familiar to the statistical psychologist from their appearance in the modifications of the normal probability curve to express skew distributions (cf. E. T. Whittaker and G. Robinson, *The Calculus of Observations*, 1924, p. 174, and Arne Fisher, *Frequency Curves*, 1922, p. 31) and in Pearson's formula for the tetrachoric correlation (P. Elderton, *Frequency Curves and Correlations*, 1938, p. 247).

of particles in the beam. But, since probability depends on relative frequency, we may also say that the value thus found expresses the probability that, in a snapshot examination made at a given instant, we shall find an electron (or whatever particle we are studying) in a given small region of space at the point specified[1].

Operators. All the foregoing arguments may readily be generalized to three or even more dimensions, and to systems or assemblies containing as many particles as we please. In dealing with an atom of helium for instance, Ψ will be a function of the coordinates for each of the two electrons; and will thus imply a ' vibration ' in a space of *six* dimensions. This makes it clear that the vibrations are ' waves ', not in any actual ' medium ', but simply in a symbolic mathematical model. Accordingly, for the more general cases it is customary to adopt Jacobi's notation, and substitute q for x, y, and z, in order to indicate that we are dealing with ' generalized ' coordinates.

The manipulation of these more complicated expressions is greatly facilitated by the use of operators. Here it will be sufficient to illustrate the method by applying it to the simplest possible case. It will be remembered that solutions to equations like eqn. (1) may be expressed by a sine, a cosine, or a combination of both; and the latter can in turn be expressed very simply by an exponential function containing ' imaginary ' or ' complex ' quantities. Let us then rewrite eqn. (14) in the form

$$\Psi = A \exp\{2\pi i(q/\lambda - \nu t)\}. \qquad \ldots (25)$$

Differentiating with respect to q, we have

$$\frac{d\Psi}{dq} = 2\pi i\Psi/\lambda \quad \text{and} \quad \frac{d^2\Psi}{dq^2} = -4\pi^2\Psi/\lambda^2.$$

But, as we have already seen (eqn. (15)), $\lambda = h/p$ and $\lambda^2 = h^2/2m(E\nu) = h^2/2mH$ say. Hence we may write the foregoing

$$\left(\frac{h}{2\pi i}\frac{d}{dq}\right)\Psi = p\Psi \quad \text{and} \quad \left(-\frac{h^2}{8\pi^2 m}\frac{d^2}{dq^2}\right)\Psi = H\Psi.$$

Thus p and H may both be represented by the differential operators placed in brackets.

Now Schrödinger's starting point (first paper, *loc. cit. sup.*, p. 1) is the Hamiltonian equation for energy, which expresses the total energy, E, as a function solely of the generalized coordinates and the corresponding moments (the so called ' Hamiltonian function '), i.e.

$$E = H(p, q) = p^2/2m + U(q). \qquad \ldots (26)$$

Here too he substitutes for p the operator $(h/2\pi i)(d/dq)$, and then applies the whole expression to the function ψ. We thus obtain for the more general case

$$\left\{\frac{d^2}{dq^2} + \frac{8\pi^2 m}{h^2}(E-U)\right\}\psi = 0, \qquad \ldots (19 \text{ bis})$$

which gives us at once the equation we have already reached by a more familiar but more circuitous approach, namely, eqn. (19). And generally it may be said that the regular procedure in ' wave mechanics ' in its simplest form is to start with the classical equations, substitute operators for observables (e.g. for the coordinates, the linear and the angular momenta, the kinetic energies, etc.); and then solve the modified equations so as to obtain Eigen-values and Eigen-functions.

The Determination of Averages and Transition Probabilities. Two of the more important results obtained by this generalized procedure may be briefly referred to here, largely to note the formal similarities between the resulting expressions and those already familiar to the factorist.

[1] Schrödinger himself originally assumed that $|\psi|^2$ represented the actual density of the charge on the electron. But this interpretation ceases to be tenable when the Schrödinger wave is no longer confined within an atom. It was to meet this predicament that Max Born suggested that " the intensity of the wave should be regarded as indicating not charge density, but probability density " (*Atomic Physics*, pp. 156f.).

When a large number of measurements are made of an observable for an assembly of systems whose state function is Ψ, then the average value may be obtained by applying the formula

$$\bar{a} = \int \Psi a \Psi^* \, dq \Big/ \int \Psi \Psi^* \, dq, \qquad \ldots (27)$$

where a denotes the operator corresponding to the observable and dq stands for a volume element in the multi-dimensional space (e.g. dx_1, dy_1, dz_1, dx_2 ... dz_n if we are dealing with a system containing n particles in three-dimensional space). The purpose of the denominator is merely to normalize the total distribution. The statistical psychologist will note the analogy between this procedure and the method of ' weighted summation ' in factor analysis —an operator, as usual, taking the place of a matrix.

To investigate the probability of a transition from one state of energy to another, it is necessary to calculate the integral

$$T_{mn} = \frac{1}{N_m N_n} \int_{-\infty}^{+\infty} \phi_n \phi_m q \, dq \qquad \ldots (28)$$

where n and m refer to the two different states whose energy levels are E_n and E_m respectively (as deduced from eqn. (23)). N denotes a normalizing factor, and ϕ the Eigen-function. It will be observed that the expression on the right has the form of a product-sum, not unlike that used for calculating covariances or correlations. In the case of a linear harmonic oscillator it turns out that the transitions can only take place between adjacent states, for which $m = n \pm 1$, and T_{mn} then reduces to $\sqrt{\{(n+1)/2\}}$ or $\sqrt{(n/2)}$.

The Principle of Indeterminacy. One final question still remains. In classical mechanics $pq = qp$, so that $pq - qp = 0$. If, however, we substitute operators, we obtain

$$pq - qp = (h/2\pi i) \frac{d}{dq}(q) - q(h/2\pi i) \frac{Ld}{dq}(1) = h/2\pi i, \qquad \ldots (29)$$

which is Heisenberg's equation demonstrating the ' indeterminacy ' of the ' canonically conjugated ' variables p and q.

The Probabilistic Interpretation. Most psychologists will doubtless be interested, not in the formal mathematical models adopted in the theory of wave mechanics, but rather in their interpretation. During earlier discussions of the ' new ' quantum theory the majority of the contributors, as we have seen, strongly deprecated any kind of concrete interpretation. Today, however, nearly every leading physicist—including Heisenberg, Born, Bohr, Dirac, and Pauli—interprets the basic formulae as expressing probabilities[1]. Although the adoption of such a view plainly implies a repudiation of determinism, it does not necessarily entail a rejection of causality in the broader sense of rational dependence. But, as Born observes, " the observable aspects of reality for which such a dependence is claimed are no longer those of classical physics: they are not the elementary events themselves, but their probabilities. We thus have the paradox that events obey the laws of chance, but their probability obeys laws which in all essential features are causal laws ", i.e., laws of dependence such as will provide a rational basis for verifiable predictions (*Natural Philosophy of Cause and Chance*, 1949, pp. 102–3). In short, " probability is fundamental in every statement about reality " (Reichenbach, *op. cit.*, p. 10).

[1] The most effective demonstrations are those of Max Born. Broadly speaking, his procedure is to derive the standard functions for ' states '. ' waves ', ' transitions ', etc., and then to show that certain coefficients appearing in these functions have the formal character of expressions describing probabilities. The same conclusion is reached by somewhat different proofs by Reichenbach (*Philosophic Foundations of Quantum Mechanics*, 1944). This mode of explaining the dualism of wave and particle seems to have started with Einstein's tentative interpretation of the similar dualism in the case of light: light waves he describes as " a ghost field, which has no ordinary physical meaning, but whose intensity determines the probability of the appearance of photons ". De Broglie suggested the alternative theory of the ' pilot ' wave guiding the particle; but, as he has since pointed out, the two interpretations are by no means incompatible.

BOOK REVIEWS

Admission to Grammar Schools. By ALFRED YATES and D. A. PIDGEON. London: Newnes Educational Publishing Co., 1957. Pp. xvi+260. 25*s*.

This is the ' third interim report ' in the series published by the National Foundation for Educational Research and dealing with the allocation of ' primary school leavers ' to courses of secondary education. It " represents the climax, though it may not mark the end, of the large-scale programme of research launched in 1951 "; and its central problem is the socalled " eleven plus examination ". The opening chapters deal with current methods of allocation, and summarize previous inquiries on the problem. The plan adopted for the present research was twofold—first, to carry out " an overall survey of the procedures used by education authorities for effecting transfer ", and, secondly, to conduct " a detailed follow-up during the secondary school careers of two consecutive age-groups in the area of one authority "—a borough in Middlesex.

The writers state that, of the various batteries of tests and assessments they have examined, "a number are able to provide validity coefficients of 0·90 or slightly more". Over the whole of the country, however, the average validity is almost certainly lower; and for this they suggest a rough estimate of 0·85—which, as they admit, may possibly err on the generous side. Considering the difficulties that inevitably attend any practicable scheme of selection, and remembering that during post-pubertal years the interest, industry, and rate of maturation of individual pupils are liable to fluctuate appreciably, they conclude that " even the best methods are still likely to involve wrong allocations for about 10 per cent. of the candidates ". In three respects, it is suggested, there is room for improvement: first, " the disturbingly large variation in grammar school provision from one area to another "; secondly, " the use made of teachers' estimates of children's potentialities "; and thirdly " the determination and treatment of the border zone ". However, they are inclined to believe that " few border-zone procedures are likely to add much to the effectiveness of methods which already yield validity coefficients in the neighbourhood of 0·90 ".

The figure on which most reviewers in the popular journals have seized is the proportion of ' wrong allocations '. " In 1955 ", says the report, " approximately 640,000 were eligible for transfer to secondary schools, and our estimates suggest that approximately 78,000 were wrongly allocated " (p. 145). The writers themselves, however, emphasize the limitations to be attached to this broad statement. And in the heading to the section that leads up to this conclusion they put the word ' wrong ' between inverted commas: it means ' wrong ' as judged by their criterion. This was a follow-up procedure which was (i) carried only as far as ages 13 and 14 and (ii) based on tests designed so as to be applicable to pupils in modern schools as well as those in grammar schools. Within the different schools marks and orders of merit were obtained from the teachers; but, for the final comparison, these had necessarily to be " quantified and scaled " by tests of " general ability and attainment in English and arithmetic " suitable for *every* child. Thus a child transferred to a grammar school is said to have been ' wrongly ' allocated if at about 13 or 14 he did not do so well as others who had been transferred to a modern school in tests *suited for pupils of the modern school type*. But bright pupils who can shine at algebra and geometry may do far worse in arithmetic than a duller child who has been taught in a modern school. It follows, therefore, as the authors point out, that " the 38,500 who remained in grammar schools, although *by our criteria they may be regarded as ' wrongly allocated ', may nevertheless have derived benefit* " (our italics).

Thus it is quite misleading to treat their estimates (in the words of one reviewer) as a " grave indictment of the selection procedures actually adopted ". Such a charge can only be sustained if the critic is able to suggest a *better practical procedure* which would diminish the number of ' wrong allocations '. The fact of the matter is, as Pidgeon and Yates observe, those who attack the selection procedure are really criticizing it for not performing a ' virtually impossible task ',

The practical corollaries would seem to be, first, that the training and opportunities available for the brightest of those transferred to modern schools should not diverge too widely from what is provided for the duller pupils transferred to the grammar schools; secondly—and perhaps most important of all—there should be a more sustained and effective provision for checking subsequent progress and far greater facilities for subsequent cross-transference. From the figures given in the report it would appear that, between the ages of 12 and 15, only about 1 in 200 of those ' wrongly ' allocated to modern schools find their way sooner or later to a grammar school. From this it might of course be argued that most of the socalled ' wrong ' allocations were not wrong in the view of the teachers. However, the figure for those subsequently rejected after one or two years at the grammar school (0·5 per cent.) is surprisingly low. The proportion of students rejected after one or two years at a university (where the initial selection is far more rigorous) amounts on an average to well over 7 per cent.

The results of the ' follow-up ' provide ample evidence in favour of the retention of the socalled ' intelligence ' tests. This conclusion seems further supported by the practical experience of education authorities themselves. Since 1952 the number employing standardized tests of this type has markedly increased; among County Councils it has risen from 82 per cent. to 91 per cent. Of the different methods of assessment used in the present research, the head teachers' ratings appeared on the whole to be the best; a verbal ' intelligence ' test almost as good; the results of the attainments tests were much poorer, and those of the non-verbal or ' spatial ' tests poorest of all. It would seem, however, that the kinds of problem used are still, in all essentials, the same as those introduced by Burt for similar purposes nearly fifty years ago, and the validity coefficients of the constituent subtests are no higher than those achieved after the first four or five years' trial. Surely by now, with all the improvements in factorial techniques and computing procedures, it should be possible for some equally ingenious or enterprising research worker to vary and improve the basic type of test.

As to the effects of environment, the writers' conclusions are at present merely tentative, since this was not a question with which they were primarily concerned. So far as their own evidence goes, they are inclined to endorse the conclusion of previous investigators that " a child's chances of success in a grammar school are affected by the extent to which he receives support and encouragement from his parents ", but they reject the suggestion that present methods of selection are unjust to what are still called " the unprivileged classes ". Even the widespread notion that a verbal type of intelligence test favours " children from higher social groups " is not, they say, supported by their own experience. But here perhaps their arguments would carry still greater conviction if (as has so often been urged) the ambiguous phrase ' intelligence test ' were abandoned, and some such designation as ' general classification test ' adopted instead. As was pointed out when group tests of this type were first introduced, they do not measure ' pure intelligence ' in the sense of ' innate potentiality '. But for the practical purposes of allocation this is no longer the sole criterion.

Not unnaturally parents whose children fail to secure a place in the grammar school feel dissatisfied; and, as the report points out, their dissatisfaction is commonly vented against the novel features in the techniques used. Such complaints, it is contended, are " wholly misplaced ": the real object of criticism should rather be " the lack of a generally accepted policy and the unequal provision of grammar school places ".

The report as a whole presents us with the most intensive study hitherto made of the problems raised. It concludes with a number of detailed appendices on statistical issues by Dr. P. M. Grundy, which will be well worth perusal by every psychologist or educationist who contemplates a further investigation of this type. There are minor points both of method and of presentation which might easily be criticized. But, when we bear in mind the practical limitations under which all theoretical inquiries of this type necessarily labour, the severest critic can have little but praise for the way the investigation has been conducted and for the cooperation received from the authorities and the teachers. A. G. H.

Statistical Theory: The Relationship of Probability, Credibility, and Error. By LANCELOT HOGBEN. London: Allen & Unwin, 1957. Pp. 510. 45*s*.

On the title page Dr. Hogben describes this work as " an examination of the contemporary crisis in statistical theory from a behaviourist viewpoint ". He believes that the studies of J. Neyman, E. S. Pearson and A. Wald " threaten to undermine the entire superstructure of estimation and test procedure erected by R. A. Fisher on foundations laid by Karl Pearson, Edgeworth, and Yule "; and his book is largely an attempt, not altogether convincing, to carry on the attack by further arguments of his own.

To the psychological reader the chapter of greatest interest is that entitled " The Impasse of Factor Analysis ". The methods adopted by the psychological factorist are, as Dr. Hogben points out, due to Karl Pearson; and in his view they suffer from the fundamental defects which he believes to have been inherent in the ' Pearsonian ideology '. The procedure, however, which he expounds in greatest algebraic detail is that put forward by Spearman as a simplification of Pearson's method. The views of Hotelling, Thurstone, and Thomson are discussed more briefly. He contrasts Burt's view of the ' law of parsimony ' with that expressed by Thurstone, and adds that " the controversy between the different schools as revealed in the passage quoted lays bare an issue, raised elsewhere ", namely, " that there are as many criteria of economically interpreting nature as there are different frameworks of interpretation ". His main conclusion most British factorists today would readily endorse. " Factor analysis ", he says, " like multivariate analysis, is merely descriptive. At best it can lead to a more satisfactory taxonomy, but only if we are clear about what and why we want to classify: it cannot disclose a causal nexus ". It is curious that he does not recognize that nearly all these criticisms have already been put forward by Burt and others in this country.

Most of the chapters seem far more hastily written, and consequently far more difficult to follow, than his previous publications on statistical problems. Nevertheless throughout the work the critical reader will find much that is at once suggestive and provocative.

A. K. S.

British Philosophy in the Mid-Century. Edited by C. A. MACE. London: Allen & Unwin, 1957. Pp. 396. 30*s*.

The essays published in this volume originated in a series of lectures organized by the British Council at Cambridge in 1953. The contributors include twelve of the most eminent philosophers in the country. The general psychologist will doubtless turn first to the papers by Ayer on ' Perception ', Moore on ' Visual Sense Data ', and Mace on ' Trends in the Philosophy of Mind '—all brilliantly written. The statistical reader will find Braithwaite's discussion of ' Probability and Induction ' and Popper's appendix on the same problem full of instructive points, even though he may not agree with their somewhat controversial conclusions. But in many ways the most interesting paper of the whole series is the introductory survey by Professor Broad of ' The Logical Historical Background of Cambridge Philosophy '. Here he reviews the work of fourteen Cambridge philosophers, most of whom Broad himself knew personally. He notes how Ward's work " created a revolution in English psychology "; and remarks that, although the views which Ward put forward are at the moment " highly unpopular at present in England and the U.S.A., this does not preclude the possibility that Ward may be substantially right and the fashionable opinion fundamentally mistaken ". Of the contributions made by the others in his list, which includes Stout, Moore, Whitehead, Neville and Maynard Keynes, and Lord Russell, we are given an admirably succinct and illuminating study, at once sympathetic and critical.

T. F.